

# Neural Mechanisms Underlying Trust to Friends, Community Members, and Unknown Peers in Adolescence

Sophie W. Sweijen<sup>1,2</sup> , Suzanne van de Groep<sup>1</sup>, Lysanne W. te Brinke<sup>1</sup>,  
Andrew J. Fuligni<sup>3</sup>, and Eveline A. Crone<sup>1,2</sup>

## Abstract

■ Trust plays an important role during adolescence for developing social relations. Although prior developmental studies give us insight into adolescents' development of differentiation between close (e.g., friends) and unknown (e.g., unknown peers) targets in trust choices, less is known about the development of trust to societal targets (e.g., members of a community organization) and its underlying neural mechanisms. Using a modified version of the Trust Game, our preregistered fMRI study examined the underlying neural mechanisms of trust to close (friend), societal (community member), and unknown others (unknown peer) during adolescence in 106 participants (aged 12–23 years). Adolescents showed most trust to friends, less trust to community members, and the least trust to

unknown peers. Neural results show that target differentiation in adolescents' trust behavior is associated with activity in social brain regions implicated during mentalizing, reward processing, and cognitive control. Recruitment of the medial prefrontal cortex (mPFC) and OFC was higher for closer targets (i.e., friend and community member). For the mPFC, this effect was most pronounced during no trust choices. Trust to friends was additionally associated with increased activity in the precuneus and bilateral temporal parietal junction. In contrast, bilateral dorsolateral prefrontal cortex and anterior cingulate cortex were most active for trust to unknown peers. The mPFC showed increased activity with age and consistent relations with individual differences in feeling needed/useful. ■

## INTRODUCTION

Adolescence is an important transition period between childhood and adulthood, in which individuals develop perspective taking and social competencies that are needed to eventually develop into contributing members of society (Fuligni, 2019, 2020; Crone & Dahl, 2012). One of the fundamental tasks during adolescence is to develop mature social relationships and societal values (Crone & Fuligni, 2020). Trust is one of the crucial building blocks for successfully developing social relationships (Crone, Sweijen, te Brinke, & van de Groep, 2022; Burke, van de Groep, Brandner, & Crone, 2020; Crone & Dahl, 2012; Lahno, 1995). Trust requires advanced levels of perspective taking and strategic thinking, which are processes that show continued development across adolescence (Dumontheil, Apperly, & Blakemore, 2010). We recently discovered that adolescents trust institutional community members less than friends but more than unknown peers (Sweijen, te Brinke, van de Groep, & Crone, 2023), suggesting that adolescents show intermediate trust to more distant but potentially personally relevant institutions. These findings fit with recent studies showing developmental

differences in prosociality toward close and distant others across adolescence (Sweijen et al., 2023; Fett, Shergill et al., 2014; Güroğlu, van den Bos, & Crone, 2014). Although previous studies suggest that adolescent relationships with peers mostly rely on interpersonal trust (Güroğlu, 2021), it remains unknown whether adolescent relations with other societal targets (e.g., community members) rely on trust similarly. Given that adolescents discover their position in society and develop these broader connections (Fuligni, 2019), understanding the underlying processes of trust, particularly community-based trust, is crucial.

Functional neuroimaging studies can shed light on the underlying processes of trust choices and whether this is dependent on the target to which trust is shown. The Trust Game is an experimental economic exchange paradigm in which an individual (the trustor) can trust a certain number of resources (e.g., coins) to another individual (the trustee), after which those resources may increase depending upon the decision of the trustor. The trustee then has the power to either reciprocate (share the resources) or defect (keep resources for self). Some developmental studies found age-related increases in trust between childhood and adolescence (van den Bos, van Dijk, Westenberg, Rombouts, & Crone, 2011; van den Bos, Westenberg, van Dijk, & Crone, 2010; Sutter & Kocher, 2007), whereas other studies showed a general stability in trust (van de Groep, Meuwese, Zanolie,

<sup>1</sup>Erasmus School of Social and Behavioral Sciences, Erasmus University Rotterdam, The Netherlands, <sup>2</sup>Department of Developmental and Educational Psychology, Institute of Psychology, Leiden University, The Netherlands, <sup>3</sup>University of California

Güroğlu, & Crone, 2018; Lemmers-Jansen, Krabbendam, Veltman, & Fett, 2017; Fett, Shergill et al., 2014; Güroğlu et al., 2014). The exact developmental patterns of trust, particularly during adolescence, seem dependent on social contextual factors, such as the target of trust (Güroğlu et al., 2014), the risk for the trustor (van de Groep et al., 2018), and the behavior of the trustee in a multiple-round Trust Game or in daily life (Fett, Shergill et al., 2014; Güroğlu et al., 2014).

Given the sociocognitive strategic demands needed in the interaction between the trustee and trustor (Rilling & Sanfey, 2011), trust was previously found to be associated with flexible recruitment of brain areas underlying sociocognitive developmental processes (Crone et al., 2022; Burke et al., 2020), including brain regions associated with mentalizing, reward processing, and cognitive control. Mentalizing, which is defined as the ability to understand the mind of one's own and other individual's, has previously been associated with activation in social brain regions (Blakemore & Mills, 2014; Frith & Frith, 2012; Lieberman, 2007), including the medial prefrontal cortex (mPFC), TPJ, and precuneus (Crone & Dahl, 2012; Blakemore, 2008). However, although these regions are mostly associated with social cognition, recent work suggests that "social brain" regions encompass both social and nonsocial processes (Konovalov, Hill, Daunizeau, & Ruff, 2021; Lockwood, Apps, & Chang, 2020). Nevertheless, studies using the Trust Game showed that young adults engage the mPFC during trust choices (Cutler & Campbell-Meiklejohn, 2019; van den Bos, van Dijk, Westenberg, Rombouts, & Crone, 2009), although this is not consistently found in younger adolescents (Sijtsma et al., 2023). Moreover, previous developmental studies showed age-related increases in activity in these regions, possibly indicating advancing mentalizing skills across adolescence (Lemmers-Jansen et al., 2017; Fett, Gromann, Giampietro, Shergill, & Krabbendam, 2014; Fett, Shergill et al., 2014; van den Bos et al., 2011), but how these brain regions typically associated with mentalizing are engaged when giving trust to community members relative to friends and unknown others remains unknown.

In addition to the role of the mentalizing brain regions during trust (Blakemore & Mills, 2014; Frith & Frith, 2012; Lieberman, 2007), reward processing and cognitive control are also expected to play an important role during trust choices, specifically to close others. That is, brain regions associated with reward processing, including the OFC and ventral striatum (i.e., caudate, putamen, and nucleus accumbens [NAcc]; Lemmers-Jansen et al., 2017; Fett, Gromann et al., 2014; van den Bos et al., 2009; Delgado, Frank, & Phelps, 2005), is involved in an individual's motivation to cooperate with others (Declerck, Boone, & Emonds, 2013; King-Casas et al., 2005). For example, Fett, Gromann, and colleagues (2014) found increased activity in social reward-related areas, specifically in the OFC, during trust choices. Age was positively associated with the posterior cingulate/precuneus and negatively associated

with the OFC, left and right caudate nucleus, and bilateral PFC, possibly reflecting an increased sensitivity toward social signals with increasing age (see also Lemmers-Jansen, Fett, Shergill, van Kesteren, & Krabbendam, 2019). Finally, cognitive control may play an important role in the strategic motivation to cooperate (Declerck et al., 2013), such that individuals rely on brain regions associated with cognitive control while regulating self-oriented impulses involved in trust choices. Prior studies reported age-related increases in the ACC and dorsolateral PFC (dlPFC) related to trust (Lemmers-Jansen et al., 2017; Fett, Gromann et al., 2014; van den Bos et al., 2009), but no prior studies examined engagement of these regions in the context of giving trust to community members, although this may strongly depend on strategic motivations as well as adolescent-specific transitions in their roles in a larger societal context.

Some previous studies examined how brain regions involved in mentalizing, reward processing, and cognitive control show target differentiation for prosocial giving tasks. These studies found that prosocial choices involving liked compared with neutral or disliked others, determined via sociometric nominations, were associated with increased activity in NAcc, TPJ-intraparietal lobe (TPJ-IPL), and ventromedial PFC for liked others (Schreuders, Klapwijk, Will, & Güroğlu, 2018; Güroğlu et al., 2008). Furthermore, Telzer, Masten, Berkman, Lieberman, and Fuligni (2011) demonstrated increased activity in dlPFC, dorsomedial PFC, and precuneus, regions involved in cognitive control and mentalizing, during prosocial decisions about specific others who are viewed as similar or close to oneself, such as family members. In addition, Karan and colleagues (2022) showed that the dlPFC and the ventrolateral PFC showed age-related increases as adolescents increasingly differentiated in their giving behavior between targets. Finally, prior developmental studies using giving paradigms demonstrated increased activity in adolescents in the anterior IPL, TPJ, putamen, and right lateral PFC (lPFC) during giving to friends compared with unfamiliar peers, as well as age-related increases in activity in the lPFC when giving to unfamiliar peers (van de Groep et al., 2022; van de Groep, Zanolie, & Crone, 2020; Schreuders, Smeeckens, Cillessen, & Güroğlu, 2019; Schreuders et al., 2018). The authors suggest that this increased neural activity to friends relative to unknown peers possibly reflects increased perspective-taking, often referred to as cognitive empathy (van den Bos et al., 2010).

Taken together, prior studies using giving paradigms suggest distinct neural correlates of prosocial decisions to close and distant targets, which gives a first indication that mentalizing brain regions may respond differentially when giving trust to members of a community organization. Moreover, previous studies in which participants played Trust Games with trustworthy and untrustworthy counterparts, demonstrated age-related decreases in activity in the OFC and caudate (Fett, Gromann et al., 2014) and in the anterior mPFC (van den Bos et al.,

2011) during interactions with trustworthy individuals. Given that a key developmental challenge of adolescence is to develop into a contributing member of society (Fuligni, 2019), it is important to investigate adolescents' trust to not only personally close targets (e.g., friends) but also to societal targets (e.g., community members) as well as distant or unknown others (Sweijen et al., 2023). The aim of this study was to investigate these neural activations underlying trust choices to societal versus close and unknown others. We aimed to study whether similar or distinct neural regions underly trust toward community members, with whom individuals have no interpersonal relationship.

## The Current Study

The current study, including our hypotheses, was preregistered on the Open Science Framework (see <https://osf.io/f245>). Using a modified version of the Trust Game (Sweijen et al., 2023; Güroğlu et al., 2014; Berg, Dickhaut, & McCabe, 1995), the aim of our study was to examine the underlying neural mechanisms of trust to close, societal, and unknown others (i.e., friends, community members, and unknown peers) during adolescence. Specifically, we examined whether trust is associated with neural activity in social brain regions implicated in mentalizing (specifically, mPFC, bilateral TPJ, precuneus), reward (specifically, NAcc and OFC), and cognitive control (specifically, ACC and bilateral dlPFC).

On a behavioral level, we expected that adolescents show more trust to community members and most trust to friends, compared with unknown peers (Hypothesis 1.1; main effect target; Sweijen et al., 2023). In a pilot study on adolescents' trust and reciprocity (Sweijen et al., 2023), we observed that general contributions to society and interpersonal trust beliefs were positively associated with trust and reciprocity. Therefore, we examined whether age is related to trust percentages to the different targets (Hypothesis 1.2; Sweijen et al., 2023; Fett, Shergill et al., 2014; Güroğlu et al., 2014), as well as to advancing developmental processes throughout adolescence, namely, perspective taking, cognitive control, feeling needed and useful, and social reward sensitivity (Hypothesis 1.3; Sweijen et al., 2023; Crone et al., 2022; Burke et al., 2020; Crone & Fuligni, 2020; Fuligni, 2019). We explored whether individual differences in these developmental processes explain the most variance in trust to community members, compared with friends and unknown peers (Crone & Fuligni, 2020). Regarding feeling needed and useful, we were interested in whether these feelings contribute to trust choices, particularly to community members as an understudied but important target, because feeling needed and useful may tap into a type of instrumental connection with others resulting in a greater likelihood of providing resources to others (Fuligni, 2019).

To test the neural correlates of trust, we first examined target effects. We expected that deliberating trust is

associated with activity in at least one of the following brain regions implicated in sociocognitive processes: mentalizing (mPFC, bilateral TPJ, and precuneus), reward (NAcc and OFC), and cognitive control (ACC and bilateral dlPFC; Crone & Fuligni, 2020; van de Groep et al., 2020, 2022; Schreuders et al., 2018, 2019; Lemmers-Jansen et al., 2017; Fett, Gromann et al., 2014; Fett, Shergill et al., 2014; van den Bos et al., 2009, 2011). In addition, we expected that neural activity in these brain regions when deliberating trust choices is influenced by the relationship with the target (Hypothesis 2.1; Schreuders et al., 2018; Telzer et al., 2011; Güroğlu et al., 2008). We expected most neural activity when deliberating trust to friends, less to community members and least to unknown peers, mirroring prior behavioral findings (Sweijen et al., 2023). We explored whether neural activity to the different targets is also dependent on the choice to trust or not to trust each target (Hypothesis 2.2).

In prior research, age-related increases in TPJ activity when interacting with untrustworthy individuals were found to be associated with age-related decreases in trust choices (Lemmers-Jansen et al., 2017; Fett, Gromann et al., 2014). Therefore, we expected brain-behavior associations, such that neural activity related to mentalizing, reward, and cognitive control correlates to general trust choices (Hypothesis 2.3). Finally, we expected age-related increases in recruitment of these brain regions (Hypothesis 2.4; Lemmers-Jansen et al., 2017; Fett, Gromann et al., 2014; Fett, Shergill et al., 2014; van den Bos et al., 2009). We explored whether there were age effects in the developmental trajectory of neural activity associated with trust to the different targets. Finally, to examine whether activity in the different brain regions is related to the underlying sociocognitive processes, we expected that, controlling for age, neural activity would be related to individual differences in perspective taking, cognitive control, feeling needed and useful, and social reward sensitivity (Hypothesis 2.5; Lemmers-Jansen et al., 2017; Fett, Gromann et al., 2014; Fett, Shergill et al., 2014; van den Bos et al., 2009). We explored differential patterns in neural activity to the different targets.

## METHODS

### Participants

Our sample included 106 adolescents (age range = 12.11–22.55 years;  $M_{\text{age}} = 18.07$  years,  $SD_{\text{age}} = 2.66$  years; 68 female participants) who participated at the third measurement wave (T3) of a cohort-sequential longitudinal project on the development of prosocial behavior in adolescence called “Brainlinks” (T1, May to October 2018; T2, August 2019 to January 2020; T3, October 2021 to June 2022). Participants were recruited through local and online advertisements. Written informed consent was obtained from participants and, in case of minors, also from their parents at each measurement wave. Participants had

normal or corrected-to-normal vision, and no intellectual disability ( $IQ < 70$ ). All 142 adolescents who participated at T1 were invited to participate at T3, if they gave consent at T1 and/or T2 to be contacted again for a follow-up study. In total, 118 participated at T3 (age range = 12.11–22.55 years;  $M_{age} = 18.07$  years,  $SD_{age} = 2.71$  years; 74 female participants). Reasons for not participating at T3 were no longer reachable ( $n = 5$ ) or no longer able to participate because of personal circumstances ( $n = 16$ ; e.g., too busy, or unmotivated). Our behavioral sample included 106 participants, because 12 participants were unable to visit our laboratory because of personal circumstances. Next, participants were excluded from our MRI sample if they had MRI contraindications and thus performed the Trust Game outside of the scanner ( $n = 10$ ), showed excessive head movement (i.e., more than 3 mm) during the fMRI task ( $n = 1$ ), or because of mask distortions ( $n = 1$ ). As such, our final MRI sample included 94 participants (age range = 12.74–22.55 years;  $M_{age} = 18.26$  years,  $SD_{age} = 2.59$  years; 60 female participants).

Post hoc power analyses using G\*Power (Version 3.1.9.7; Faul, Erdfelder, Lang, & Buchner, 2007) revealed statistical power higher than 95% considering an alpha of .05 to obtain an effect size of .2 for the behavioral sample of  $n = 106$  and the MRI sample of  $n = 94$ . We detected no differences in age ( $p = .076$ ) and sex ( $p = .588$ ) between participants who performed the fMRI task in the scanner, outside of the scanner, and those who did not perform the fMRI task. Whereas participants were initially screened by means of a checklist to verify the absence of neurological and psychiatric disorders at T1, 10 participants reported to have one or multiple disorders diagnosed by a clinician at T3 (attention deficit hyperactivity disorder:  $n = 5$ ; anxiety disorder:  $n = 2$ ; mood disorder:  $n = 1$ ; eating disorder:

$n = 1$ ; personality disorder:  $n = 1$ ; other such as addictions, sleeping problems, and emotion dysregulation:  $n = 3$ ). In addition, note that these data were collected during the COVID-19 pandemic while governmental regulations against the spread of the virus were applied in The Netherlands, which may partly explain the drop-out of participants between T3 and the previous measurement waves of the longitudinal study.

This study was approved by the medical ethical committee of Leiden University Medical Center and the Psychology Research Ethics committee of Leiden University. Participants received €50 and small presents as compensation, alongside with additional small earnings from the fMRI task and other tasks as part of the larger study protocol.

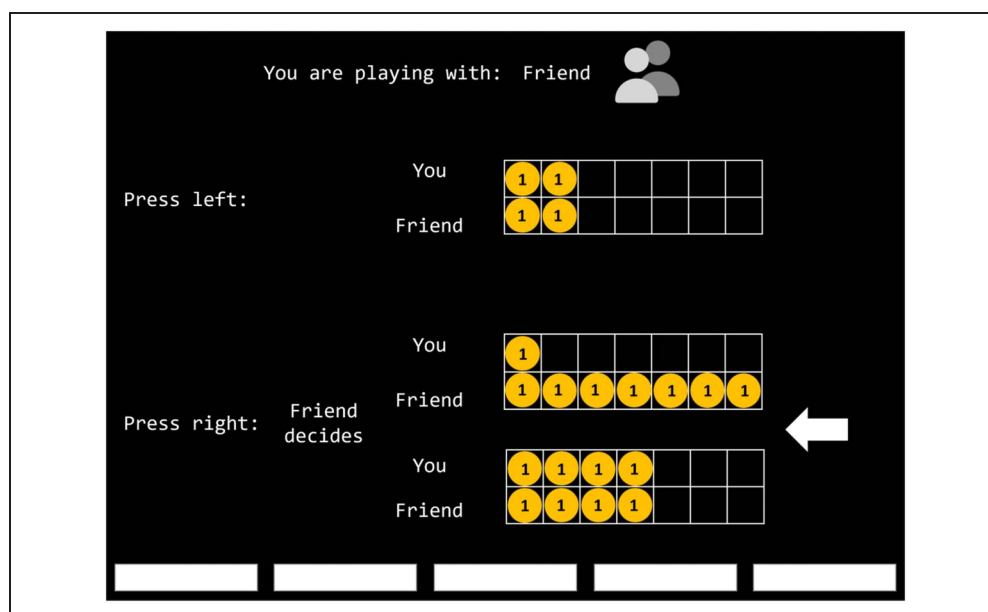
## Materials

### Societal Trust Game

Participants performed a modified version of the Trust Game in the MRI scanner (Sweijen et al., 2023; Güroğlu et al., 2014; van den Bos et al., 2009; Berg et al., 1995). In this game, participants played as the first player (the trustor) and received instructions that they would play one-shot Trust Games using fixed coin distributions with several age-matched individuals as second players (the trustees). Participants received on-screen instructions while the experimenter read aloud the instructions. Participants as the first player received 4 coins and could make two different choices (see Figure 1): either to divide the coins themselves equally (2–2) or to give coins to the second player, after which the coins in that trial were multiplied by two. In the latter case, the second player could decide how to divide the coins (1–7 or 4–4).

Participants played the Trust Games with different age-matched targets as second players, including the

**Figure 1.** Example trial of the Trust Game to the participants' friend. While the entire decision tree was shown, participants chose how to divide the coins by pressing the left or right button. The arrow indicated the response made by the participant, which was displayed throughout the trial. Blocks at the bottom of the screen indicated the time participants had left.





participant's close friend, a community member, and an unknown peer. The order in which the trials with the different targets were presented to the participants was randomized. Participants were explained that the other individuals were actual players who could also earn coins (given the importance of including real counterparts; see Johnson & Mislin, 2011) and that they would be matched with these individuals. First, participants could nominate a close friend to play the game with. To make the relation with the target comparable across participants, this could not be someone with whom they had a romantic relationship. Participants were explained that the experimenters would contact this friend by sending them a text message to invite them to play the game too, to minimize deception. To ensure this, we contacted and invited each friend after the laboratory visit to play the game (online). Second, the community member was a young individual of the Dutch National Youth Council (NJR). Participants received information about this community, namely, that this youth community aims to defend the interests of young people in The Netherlands by, for example, organizing youth panels. In addition, participants watched an informational video in which the community and their mission were introduced. We contacted the community NJR and asked one of their members to play several trials of the Trust Game as the second player for the community trials. It was ensured that the coins this community member earned during the game were donated to NJR. Third, participants were told that the unknown peer was another participant of the study of approximately the same age. An age-matched student played several trials of the Trust Game as the second player for the unknown peer trials. By using actual counter-players in our design of the Trust Game, we increased the likelihood that the decisions of participants were made based on playing with real counterparts.

Participants were instructed that they could earn actual money by playing the game for themselves and the other players, such that both players would receive coins earned during three random trials selected by the computer (each of these three trials with a different player). The choices of both players would decide the number of coins each player would receive. Although coin distributions were fixed across all trials, variation was added to the game to keep participants engaged by using different coins representing different coin values (yellow coin 10 eurocents, blue coin 20 eurocents, and orange coin 30 eurocents). Although different coin values were used to increase participants' engagement during the game, similar to the study of van Hoorn, van Dijk, Güroğlu, and Crone (2016), we averaged across these values in our analyses. In total, these coin values resulted in a potential sum of money ranging from 60 eurocents to 6 euros and 30 cents each player received. We transferred the money to all players (i.e., the participant, friend of the participant, and community member) after all players played the Trust Game.

The Trust Game was presented in the MRI scanner via E-prime Version 3. An example trial is shown in Figure 1. After a fixation cross, which was jittered from 500 to 3000 msec to better estimate the hemodynamic response function (HRF; Miezin, Maccotta, Ollinger, Petersen, & Buckner, 2000), the entire decision tree was shown. On each trial, the name (e.g., actual name of the friend, young individual for the community member, and stranger for the unknown peer) and an avatar of the second player were displayed at the top of the screen to indicate the target of the choice (i.e., trust or no trust). Participants indicated within 5000 msec whether they chose the "no trust" option (i.e., left button), resulting in the end of the trial, or the "trust" option (i.e., right button), allowing the trustee to decide the outcome of the trial. In the latter case, the coins in that trial were multiplied by two. Participants did not receive feedback on the subsequent choice by the second player. Note that the labels of "no trust" and "trust" were not visible to the participants. A timeline at the bottom of the screen, consisting of five blocks that disappeared one by one after 1000 msec (see Figure 1), indicates the time participants had left to press one of the two buttons on each trial.

On the basis of previous studies with adolescents (Sweijen et al., 2023; van de Groep et al., 2018; Lemmers-Jansen et al., 2017; Fett, Gromann et al., 2014; Fett, Shergill et al., 2014; Güroğlu et al., 2014), we expected trust choices in approximately 70% of the trials. The game therefore included 42 trials for each target, which allowed us to analyze approximately 30 trust choices for each target. This resulted in 126 trials in total (14 trials  $\times$  3 targets  $\times$  3 coin values). The game consisted of two blocks of each approximately 8 min, with a short break in between the two blocks. All combinations were randomly presented and equally distributed across trials. Before participants played the game in the MRI scanner, 18 practice trials were presented outside the scanner to familiarize participants with the game and the response options (2 trials  $\times$  3 targets  $\times$  3 coin values), and it was checked that participants understood the task.

After completing the Trust Game, participants rated on a 7-point Likert scale from 1 (*not at all*) to 7 (*very much*) how much they liked each of the three targets, how important each target was to them, and how much they believed to be helpful to each target by giving coins. These manipulation checks allowed us to examine whether participants differentiated between the three targets during the Trust Game.

### *Feeling Needed and Useful*

The Dutch translation of the Feeling Needed and Useful measure (Fulgini, Smola, & Al Salek, 2022a, 2022b) was used to measure the sense of being needed and useful. This questionnaire assesses an individual's sense of being needed and useful in multiple domains in their daily life, namely, friends, family, society, school, and work.

Participants were asked how often 25 statements described their feelings about each of these domains (five statements per domain) using a 5-point Likert scale ranging from 1 (*almost never*) to 5 (*almost always*). An example item is “I feel useful in my family.” The mean score across all domains was computed (Cronbach’s  $\alpha = .93$ ).

### *Social Reward Sensitivity*

The Dutch version of the Social Reward Questionnaire–Adolescent version (SRQ-A; Foulkes, Neumann, Roberts, McCrory, & Viding, 2017) was used to measure social reward sensitivity. This questionnaire consists of five subscales: admiration (four items measuring enjoyment of gaining positive attention; Cronbach’s  $\alpha = .76$ ), negative social potency (five items measuring enjoyment of using others; Cronbach’s  $\alpha = .70$ ), passivity (three items measuring enjoyment of giving control to others; Cronbach’s  $\alpha = .82$ ), prosocial interactions (five items measuring enjoyment of having reciprocal relationships; Cronbach’s  $\alpha = .66$ ), and sociability (three items measuring enjoyment of engaging in interactions with groups; Cronbach’s  $\alpha = .78$ ). Using a 7-point Likert scale ranging from 1 (*totally disagree*) to 7 (*totally agree*) participants rated 20 statements. An example item is “I enjoy feeling emotionally close to someone.” To obtain a complete examination of different social rewards, we computed the mean scores for each subscale.

### *Perspective Taking*

The perspective taking subscale of the Interpersonal Reactivity Index (Davis, 1983) was used to measure perspective taking. This subscale measures an individual’s tendency to spontaneously adopt the psychological viewpoint of another individual. Participants rated six items (Cronbach  $\alpha = .79$ ) using a 5-point Likert scale ranging from 0 (*does not apply to me at all*) to 4 (*completely applies to me*). An example item is “I try to look at everybody’s side of a disagreement before I make a decision.” We computed the sum score for this subscale (max score = 24).

### *Cognitive Control*

The Dutch translation of the Self Control Scale (Tangney, Baumeister, & Boone, 2008) was used to assess cognitive control. This scale measures an individual’s ability to control impulses. Participants were asked to rate 36 items (Cronbach  $\alpha = .90$ ) with a 5-point Likert scale ranging from 1 (*not at all*) to 5 (*very much*). An example item is “I often act without thinking through all the alternatives.” The mean score of all items was computed.

### **Procedure**

Participants were contacted by telephone and received information about the study. After agreeing to participate,

participants received a mail to fill in self-report surveys consisting of multiple questionnaires at home before the laboratory visit, which included the SRQ-A, Interpersonal Reactivity Index, and Self Control Scale. During the laboratory visit of approximately 3 hr, participants received instructions about the visit and underwent an MRI scan. The MRI scan consisted of two fMRI tasks, a functional scan, and a structural scan. Participants played the fMRI tasks on a laptop outside of the scanner in case of MRI contraindications. Finally, participants filled in additional questionnaires (including the Feeling Needed and Useful measure) and performed experimental tasks outside of the MRI scanner.

### **MRI Data Acquisition**

MRI scans were acquired with a standard whole-head coil using a Philips 3.0 T MR system (Philips Achieva TX). When possible, foam inserts were used at both sides of the participants’ head to prevent head motion. Stimuli were shown to participants on a screen behind the scanner, which was visible to the participants through a mirror attached to the head coil. Participants responded by pressing buttons using a button box. The total scan duration was approximately 45 min, which included a high-resolution 3-D T1-weighted scan for anatomical reference, a resting-state fMRI scan, and the fMRI Trust Game (two runs). The order of the scans was similar for each participant. Anatomical scans were obtained with repetition time = 7.9 msec, echo time (TE) = 3.5 msec,  $228 \times 178 \times 155$  slices, voxel size =  $1.1 \times 1.1 \times 1.1$  mm, field of view =  $250 \times 196 \times 170$  mm. Functional scans were acquired during two runs of 210 dynamics each, using T2\*-weighted gradient echo planar images, repetition time = 2200 msec, echo time = 30 msec, sequential acquisition: 38 slices, voxel size  $2.75 \times 2.75 \times 2.75$  mm, field of view =  $220 \times 220 \times 115$  mm. Five dummy scans were acquired before the start of the first functional scan of both runs.

### **MRI Data Analysis**

We performed MRI data analysis using SPM12 (Wellcome Department of Cognitive Neurology). Functional images were preprocessed using realignment, slice-time correction, spatial normalization using segmentation parameters, and spatial smoothing with 6-mm FWHM isotropic Gaussian kernel. Templates were based on Montreal Neurological Institute (MNI)-305 stereotaxic space. All functional scans were corrected for head motion using six parameters. We excluded one participant from further analyses based on excessive (i.e.,  $> 3$  mm) head motion, as well as one participant because of mask distortions. Information on motion parameters is shown in Table 1.

We used a general linear model in SPM12 to perform first-level individual analyses. The fMRI time series were modeled as a series of zero duration events time-locked to the stimulus onset convolved with the HRF. The

**Table 1.** Motion Parameters, Including Small Spikes (.5–3 mm)

|                                             | <i>Full MRI Sample<br/>(n = 96)</i> | <i>After Exclusion of Participants<br/>with &gt; .3-mm Motion (n = 95)</i> | <i>After Exclusion of Participants<br/>with Mask Distortions (n = 94)</i> |
|---------------------------------------------|-------------------------------------|----------------------------------------------------------------------------|---------------------------------------------------------------------------|
| Mean (SD)                                   | 0.09 (0.07)                         | 0.08 (0.07)                                                                | 0.08 (0.07)                                                               |
| Minimum                                     | 0.001                               | 0.001                                                                      | 0.001                                                                     |
| Maximum                                     | 3.454                               | 3.080                                                                      | 3.080                                                                     |
| Number of small spikes                      | 5.28 (1.26% of<br>Total volumes)    | 4.98 (1.18% of<br>Total volumes)                                           | 4.98 (1.17% of<br>Total volumes)                                          |
| Number of participants<br>with no spikes    | 51                                  | 51                                                                         | 51                                                                        |
| Number of participants<br>with small spikes | 45                                  | 44                                                                         | 43                                                                        |

modeled events included all possible combinations of target (friend, community member, and unknown peer) and choice (trust and no trust). We created two models to examine both target and choice effects while maximizing the number of trials. In the first model, we used target only, independent of participants' choice, which included all trials of all participants that maximized the number of trials in the three target conditions. In the second model, we included both target and choice, which resulted in a loss of participants who did not make trust or no-trust choices for one or more targets. The number of participants in each analysis is reported in the Results section.

In both models, all modeled events were added as regressors to the general linear model, in combination with a basic set of cosine functions that high-pass filtered the data (120-sec cutoff). Trials on which participants did not respond were modeled separately as covariate of no interest and were subsequently excluded from analyses. The least-square parameter estimates of the height of the best-fitting canonical HRF were used for each condition in pair-wise contrasts. These pairwise comparisons resulted in subject-specific contrast images, which we used in the second-level group analyses.

We performed whole-brain analyses at group level to investigate which brain regions were specifically activated during trust to different targets, in addition to confirmatory ROI analyses. We first performed a full factorial ANOVA with Target (friend, community member, and unknown peer) to examine neural activity during trust choices to the different targets. That is, we compared the "Community Member – Friend," "Community Member – Unknown Peer," and "Friend – Unknown Peer" contrasts (and the reversed contrasts). We subsequently performed another ANOVA with Target (friend, unknown peer, and community member) and Choice (trust and no trust) to examine whether this activity also depends on participants' choice to trust or not to trust. This resulted in the following contrasts that we compared: "Community Member No Trust – Friend No Trust," "Community Member No

Trust – Unknown Peer No Trust," "Community Member Trust – Community Member No Trust," "Community Member Trust – Friend Trust," "Community Member Trust – Unknown Peer Trust," "Friend No Trust – Unknown Peer No Trust," "Friend Trust – Friend No Trust," "Friend Trust – Unknown Peer Trust," and "Unknown Peer Trust – Unknown Peer No Trust" (and the reversed contrasts). All whole-brain regression analyses are available on Neurovault (<https://identifiers.org/neurovault.collection:13693>). All results were corrected using a false discovery rate (FDR) cluster-corrected threshold of  $p < .001$ .

On the basis of the confirmatory ROI analyses in our preregistration (<https://osf.io/jf245>), we created ROIs in SPM12 using the MarsBaR toolbox (Brett, Anton, Valabregue, & Poline, 2002). To ensure that our ROI analyses are in line with prior fMRI studies, we used Neurosynth to define our ROIs (<https://neurosynth.org/>; on December 12 2022). We extracted the mass center of each ROI from Neurosynth meta-analyses using the search terms "mentalizing" for the mPFC, precuneus, and bilateral TPJ; "reward" for the bilateral NAcc and OFC; and "cognitive control" for the ACC and bilateral dlPFC. We used MarsBaR to build 10-mm spheres around the mass centers extracted for each ROI, except for the subcortical ACC and NAcc for which we used 5-mm spheres. The coordinates of the mass center for mPFC were  $x = 0, y = 53, z = 27$ ; for precuneus  $x = 0, y = -55, z = 38$ ; for left TPJ  $x = -51, y = -58, z = 21$ ; for right TPJ  $x = 51, y = -57, z = 20$ ; for left NAcc  $x = -9.66, y = 12.1, z = -6.85$ ; for right NAcc  $x = 9.64, y = 12.8, z = -6.48$ ; for OFC  $x = 2, y = 48, z = -10$ ; for ACC  $x = 7, y = 20, z = 36$ ; for left dlPFC  $x = -37, y = 33, z = 26$ ; and for right dlPFC  $x = 40, y = 35, z = 24$ .

### Statistical Analyses

We followed almost all analyses steps as detailed in our preregistration (<https://osf.io/jf245>). Given the large number of results and to avoid overlapping approaches, we deviated from the preregistration regarding three points.

First, we do not report the confirmatory ROI analyses with target-specific trust choices as covariates to the repeated-measures ANOVA with Target (friend, community member, unknown peer) and Choice (trust, no trust) as within-subject variables, but instead report relations with general levels of trust to reduce the number of analyses. Second, we do not include the TPJ-IPL in our confirmatory analyses, given the high overlap with findings for the TPJ. Finally, we do not report the whole-brain analyses with trust behavior (i.e., general trust choices) and age in the article. However, these whole-brain analyses are available on NeuroVault (<https://identifiers.org/neurovault.collection:13693>). We performed assumption checks for all analyses to test the preregistered hypotheses. Skewness values of all variables were lower than 2, indicating no violations of normality. Regarding sphericity, we reported effects using the Greenhouse–Geisser correction in case of violations of sphericity. Because we identified one outlier for the SRQ-A subscale prosocial interactions and several outliers on neural variables (i.e., ACC Friend No Trust:  $n = 1$ ; bilateral dlPFC Friend No Trust:  $n = 1$ ; mPFC Unknown Peer Trust:  $n = 2$ ; mPFC Friend No Trust:  $n = 1$ ; bilateral NAcc Friend No Trust:  $n = 1$ ; NAcc Community Member No Trust:  $n = 1$ ; NAcc Unknown Peer No Trust:  $n = 1$ ; precuneus Unknown Peer Trust:  $n = 2$ ; precuneus Friend No Trust:  $n = 1$ ; and bilateral TPJ Friend No Trust:  $n = 1$ ) as indicated by values greater than three box-lengths from the edge of boxplots, we winsorized these outliers. Here, we report the winsorized results.

We used a Bonferroni correction for correlated variables to account for multiple testing, which resulted in a corrected alpha of .018 for behavioral analyses, a corrected alpha of .013 for target only analyses, and a corrected alpha of .010 for target by choice analyses (Sankoh, Huque, & Dubey, 1997; Uitenbroek, 1997). We report and interpret uncorrected effects resulting from analyses to test Hypothesis 1.3 (i.e., individual differences) and Hypotheses 2.3–2.5 (i.e., brain–behavior associations, age effects, and individual differences), because the aim of these analyses was to provide in-depth interpretation of the main effects of target and choice.

## RESULTS

### Behavioral Results

#### *Manipulation Checks*

To examine whether participants differentiated between the three targets in liking, importance, and perceived helpfulness, we performed a nonparametric Friedman test with target (friend, community member, unknown peer) for each subjective rating ( $n = 106$ ) given that the Kolmogorov–Smirnov and Shapiro–Wilk tests indicated violations of normality for all subjective ratings ( $p < .001$ ). Participants liked their friend most ( $M = 6.66$ ,  $SD = .58$ ), followed by the community member ( $M =$

$4.39$ ,  $SD = .82$ ) and unknown peer ( $M = 3.90$ ,  $SD = .70$ ),  $\chi^2(2) = 177.40$ ,  $p < .001$ . Participants also rated their friend as the most important ( $M = 6.47$ ,  $SD = .77$ ) followed by the community member ( $M = 3.65$ ,  $SD = 1.26$ ) and unknown peer ( $M = 3.01$ ,  $SD = 1.31$ ), as indicated by the Friedman test,  $\chi^2(2) = 173.26$ ,  $p < .001$ . Finally, participants believed to be most helpful by giving coins to their friend ( $M = 5.55$ ,  $SD = 1.22$ ), followed by the community member ( $M = 4.77$ ,  $SD = 1.30$ ) and unknown peer ( $M = 3.99$ ,  $SD = 1.34$ ),  $\chi^2(2) = 84.66$ ,  $p < .001$ .

### *Trust Behavior*

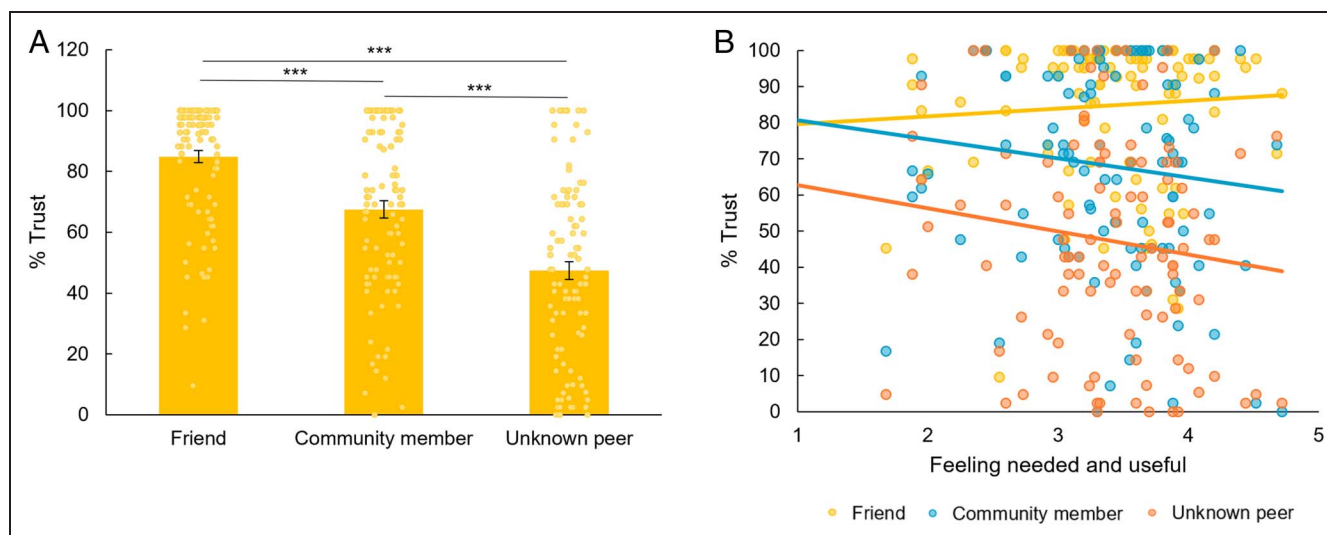
To examine whether there are differences in trust percentages to the three targets (Hypothesis 1.1), we performed a repeated-measures ANOVA with Target (friend, community member, unknown peer) as the within-subject variable and Trust Percentages as the dependent variable ( $n = 106$ ). We found a main effect of Target,  $F(2, 210) = 82.16$ ,  $p < .001$ ,  $\eta_p^2 = .44$ . As shown in Figure 2A, participants showed most trust to their friend ( $M = 84.84$ ,  $SD = 19.99$ ), less trust to the community member ( $M = 67.48$ ,  $SD = 29.00$ ), and the least trust to the unknown peer ( $M = 47.42$ ,  $SD = 30.10$ ). Descriptive statistics and correlation matrices can be found in Tables 2 and 3. Adding sex as between-subjects variable to examine sex differences in trust percentages resulted in no interaction between sex and trust ( $p = .440$ ).

Given that we averaged across the different coin values in our analyses, we additionally explored potential differences in trust behavior as a function of coin values by performing repeated-measures ANOVA with Target and Coin Value (1, 2, 3) as within-subject variables and Trust Percentages as the dependent variable. In addition to a main effect of Target,  $F(2, 210) = 82.10$ ,  $p < .001$ ,  $\eta_p^2 = .44$ , this analysis also resulted in a main effect of Value,  $F(1.41, 147.64) = 43.17$ ,  $p < .001$ ,  $\eta_p^2 = .29$ , as well as an interaction between Target and Value,  $F(3.55, 373.07) = 24.73$ ,  $p < .001$ ,  $\eta_p^2 = .19$ . Although coin value did not affect trust percentages to friends ( $p = .324$ ), coin value affected participants' trust percentages to the community member and unknown peer (both  $p < .001$ ), such that trust to these targets decreases as the coin value increases.

### *Age Effects on Trust Behavior*

To test whether trust percentages were related to age (Hypothesis 1.2), we added Linear and Quadratic Age as covariates to the repeated-measures ANOVA in a stepwise manner ( $n = 106$ ). These analyses resulted in no main effects of Linear Age ( $p = .128$ ) and Quadratic Age ( $p = .147$ ), as well as no interactions between Target and Age ( $p = .124$  and  $p = .813$  for linear and quadratic age, respectively).





**Figure 2.** (A) Mean trust percentages to friend, community member, and unknown peer. \*\*\* $p < .001$ . (B) Association between the sense of being needed and useful and trust percentages, such that individuals with a higher sense of being needed and useful differentiate more between friends and unknown peers in trust choices (uncorrected correlation,  $p = .040$ ).

### Individual Differences in Trust Behavior

To examine whether trust percentages were positively associated with individual differences in developmental processes throughout adolescence (Hypothesis 1.3), we performed a repeated-measures ANOVA with Target as the within-subject variable and all individual difference measures (perspective taking, cognitive control, feeling needed and useful, and social reward sensitivity) as covariates, controlling for age ( $n = 98$ ). This analysis resulted in an uncorrected interaction between Target and Feeling Needed and Useful,  $F(2, 176) = 3.27$ ,  $p = .040$ ,  $\eta_p^2 = .04$ . As shown in Figure 2B, adolescents with a

higher sense of being needed and useful differentiated more between friends and unknown peers in trust choices. There were no other individual differences effects. Descriptive statistics and correlations among all individual difference measures are shown in Tables 2 and 3.

### Neural Results

#### Whole-brain Analyses

To examine neural activity related to trust behavior to different targets across the whole brain, we performed a

**Table 2.** Descriptive Statistics

| Measure                     | No. of Trials/Items | N   | Min Score | Max Score | M     | SD    |
|-----------------------------|---------------------|-----|-----------|-----------|-------|-------|
| % Trust to friend           | 42                  | 106 | 9.52      | 100.00    | 84.84 | 19.99 |
| % Trust to community member | 42                  | 106 | 0.00      | 100.00    | 67.48 | 29.00 |
| % Trust to unknown peer     | 42                  | 106 | 0.00      | 100.00    | 47.42 | 30.10 |
| Feeling needed and useful   | 25                  | 105 | 1.68      | 4.72      | 3.40  | 0.62  |
| Cognitive control           | 36                  | 100 | 2.03      | 4.32      | 3.19  | 0.54  |
| Perspective taking          | 6                   | 101 | 4         | 24        | 15.88 | 3.99  |
| Social reward sensitivity   |                     |     |           |           |       |       |
| Admiration                  | 4                   | 99  | 1.75      | 7.00      | 5.02  | 1.08  |
| Negative social potency     | 5                   | 99  | 1.00      | 5.00      | 2.01  | 0.84  |
| Passivity                   | 3                   | 99  | 1.00      | 6.33      | 3.17  | 1.35  |
| Prosocial interactions      | 5                   | 99  | 3.80      | 7.00      | 6.15  | 0.67  |
| Sociability                 | 3                   | 99  | 1.00      | 7.00      | 5.56  | 1.32  |

The minimum and maximum scores indicate the actual responses (i.e., not potential responses) to the measures.

**Table 3.** Correlation Matrices

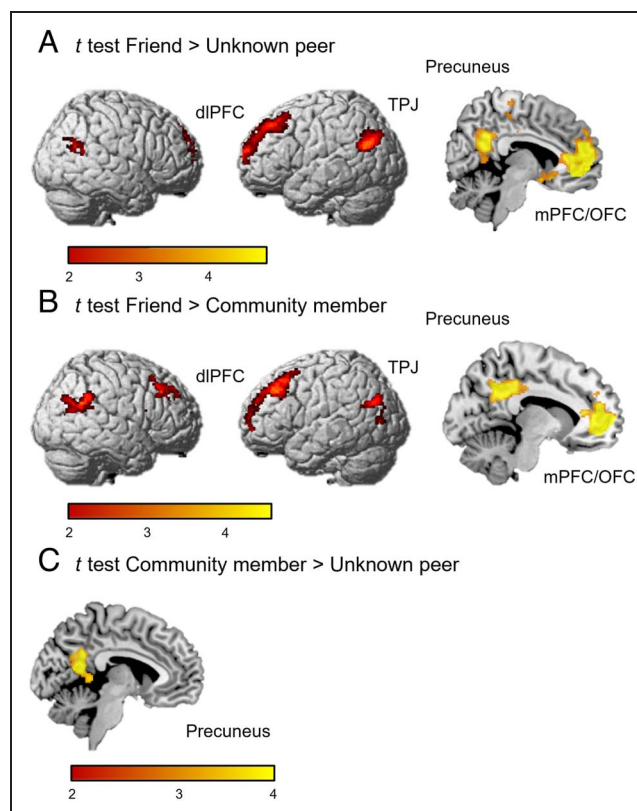
|                                           | % Trust to Friend | % Trust to Community Member | % Trust to Unknown Peer |
|-------------------------------------------|-------------------|-----------------------------|-------------------------|
| % Trust to friend ( $n = 106$ )           | —                 |                             |                         |
| % Trust to community member ( $n = 106$ ) | .36*              | —                           |                         |
| % Trust to unknown peer ( $n = 106$ )     | .22               | .51*                        | —                       |
| Age ( $n = 106$ )                         | .30*              | .10                         | .00                     |
| Perspective taking ( $n = 101$ )          | .22               | .17                         | -.03                    |
| Cognitive control ( $n = 100$ )           | -.06              | -.14                        | -.03                    |
| Feeling needed and useful ( $n = 105$ )   | .07               | -.11                        | -.13                    |
| Social reward sensitivity                 |                   |                             |                         |
| Admiration ( $n = 99$ )                   | .22               | .29*                        | .08                     |
| Negative social potency ( $n = 99$ )      | -.17              | -.03                        | -.09                    |
| Passivity ( $n = 99$ )                    | .17               | .02                         | -.12                    |
| Prosocial interactions ( $n = 99$ )       | .19               | .16                         | .06                     |
| Sociability ( $n = 99$ )                  | .12               | .24*                        | .02                     |

\*Significant at a Bonferroni corrected threshold ( $p < .018$ ).

whole-brain ANOVA with Target as well as with the interaction between Target and Choice as within-subject variable(s). All analyses are uploaded on Neurovault (see <https://identifiers.org/neurovault.collection:13693>).

**Target only.** As shown in Figure 3 and Table 4, the main effect (i.e.,  $F$  test) of Target demonstrated activation in mPFC/OFC, precuneus, bilateral TPJ and bilateral dlPFC. Paired samples  $t$  tests showed that these regions were more engaged for deliberating trust for friends compared with community members and unknown others. Only the precuneus showed significantly more activity for community members than unknown others at the whole-brain level. Together, these analyses confirmed activation in brain regions that are further examined in the confirmatory ROI analyses.

**Target  $\times$  Choice.** As indicated by an  $F$  test, there was an interaction between Target and Choice in the right insula and the right SMA (see Figure 4 and Table 5). We extracted activation for these clusters using the MarsBar toolbox. For both clusters, follow-up repeated-measures ANOVAs resulted in highly similar effects. The right insula showed more activation during no trust choices compared with trust choices,  $F(1, 57) = 7.97, p = .007, \eta_p^2 = .12$ . For both clusters, ANOVAs for the friend also showed increased activation during no trust compared with trust (both  $p < .001$ ). ANOVAs also revealed more activation for both clusters to the unknown peer during trust compared with no trust ( $p = .035$  for the right insula and  $p = .008$  for the right SMA). For trust choices, we detected increased



**Figure 3.** (A) Whole-brain  $t$  test for friend versus unknown peer. (B) Whole-brain  $t$  test for friend versus community member. (C) Whole-brain  $t$  test for community member versus unknown peer. Results are displayed with a primary voxel-wise threshold of  $p < .001$  and FDR cluster correction of  $p < .050$ .

**Table 4.** MNI Coordinates of Local Maxima Activated for the *F* Tests for Target and for the Paired-samples *t* Tests

| Area of Activation*                                                     | MNI Coordinates |          |          | Test Statistic | Cluster Size |
|-------------------------------------------------------------------------|-----------------|----------|----------|----------------|--------------|
|                                                                         | <i>x</i>        | <i>y</i> | <i>z</i> | <i>F/t</i>     |              |
| <i>F test Target (FDRc &lt; .001 = 172)</i>                             |                 |          |          |                |              |
| Frontal_Med_Orb_L                                                       | −2              | 48       | −12      | 40.16          | 5041         |
| Precuneus_L                                                             | −6              | −50      | 32       | 39.78          | 2914         |
| Caudate_L                                                               | −2              | 14       | −12      | 18.24          | 172          |
| Occipital_Mid_L                                                         | −42             | −72      | 28       | 17.58          | 588          |
| Calcarine_L                                                             | 4               | −86      | 12       | 15.34          | 331          |
| Supramarginal_R                                                         | 60              | −48      | 32       | 13.92          | 275          |
| Frontal_Inf_Tri_R                                                       | 30              | 18       | 28       | 12.14          | 405          |
| <i>t Test Friend &gt; Unknown Peer (FDRc &lt; .001 = 164)</i>           |                 |          |          |                |              |
| Frontal_Med_Orb_L                                                       | −2              | 48       | −12      | 8.89           | 5765         |
| Precuneus_L                                                             | −6              | −52      | 32       | 8.83           | 2882         |
| Occipital_Mid_L                                                         | −40             | −70      | 26       | 5.69           | 753          |
| Caudate_R                                                               | 20              | 18       | 16       | 4.63           | 187          |
| Caudate_L                                                               | −22             | 0        | 24       | 4.42           | 164          |
| Angular_R                                                               | 40              | −60      | 24       | 4.34           | 235          |
| Supp_Motor_Area_R                                                       | 4               | −26      | 68       | 4.18           | 255          |
| <i>t Test Unknown Peer &gt; Friend (FDRc &lt; .001 = 292)</i>           |                 |          |          |                |              |
| Calcarine_L                                                             | 4               | −88      | 10       | 5.05           | 292          |
| <i>t Test Friend &gt; Community Member (FDRc &lt; .001 = 316)</i>       |                 |          |          |                |              |
| Cingulum_Ant_L                                                          | −2              | 34       | 10       | 7.51           | 4740         |
| Cingulum_Mid_L                                                          | −10             | −48      | 32       | 6.39           | 2406         |
| SupraMarginal_R                                                         | 60              | −48      | 30       | 4.97           | 573          |
| Frontal_Sup_R                                                           | 22              | 34       | 32       | 4.80           | 1033         |
| Occipital_Mid_L                                                         | −42             | −74      | 28       | 4.52           | 316          |
| <i>t Test Community Member &gt; Friend (FDRc &lt; .001 = 444)</i>       |                 |          |          |                |              |
| Calcarine_L                                                             | 4               | −84      | 14       | 4.91           | 444          |
| <i>t Test Community Member &gt; Unknown Peer (FDRc &lt; .001 = 222)</i> |                 |          |          |                |              |
| Precuneus_L                                                             | −4              | −54      | 16       | 5.65           | 694          |
| Fusiform_L                                                              | −20             | −32      | −22      | 4.66           | 222          |

Results were FDR cluster corrected based on a threshold of  $p < .001$ . The other (reversed) contrasts did not result in significant effects.

\*Names were based on the AAL toolbox in SPM.

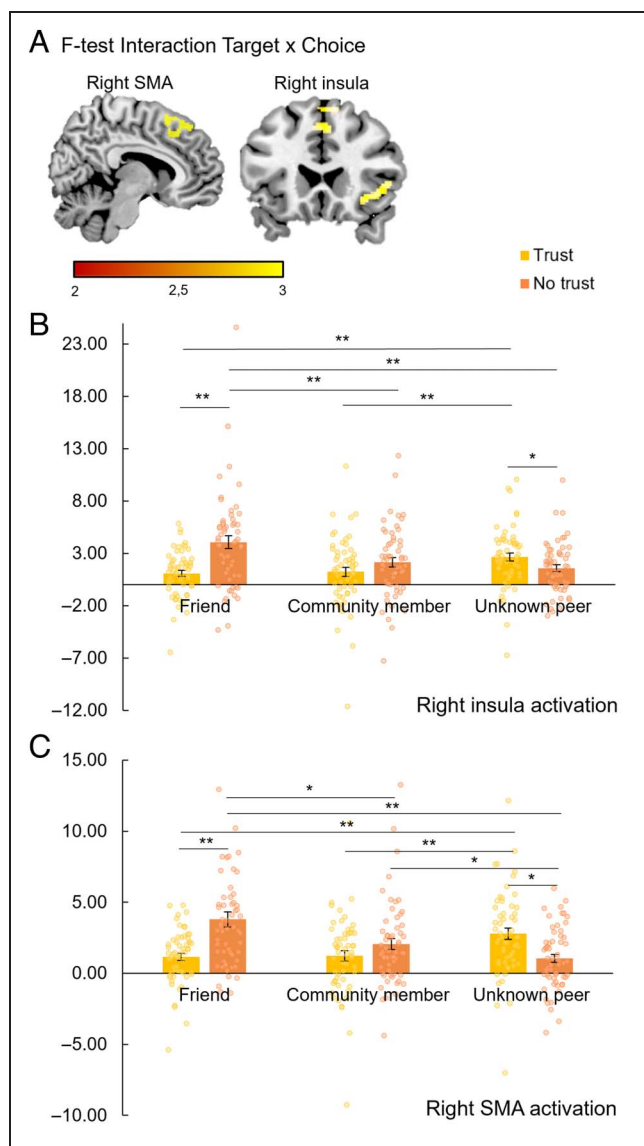
**Table 5.** MNI Coordinates of Local Maxima Activated for the *F* Tests for the Interaction between Target and Choice and for the Paired-samples *t* Tests

| Area of Activation*                                                                  | MNI Coordinates |          |          | Test Statistic | Cluster Size |
|--------------------------------------------------------------------------------------|-----------------|----------|----------|----------------|--------------|
|                                                                                      | <i>x</i>        | <i>y</i> | <i>z</i> | <i>F/t</i>     |              |
| <i>F</i> test Target × Choice (FDRc < .001 = 221)                                    |                 |          |          |                |              |
| Supp_Motor_Area_R                                                                    | 8               | 20       | 60       | 13.15          | 543          |
| Insula_R                                                                             | 32              | 20       | −10      | 10.47          | 221          |
| <i>t</i> Test No Trust > Trust (FDRc < .001 = 403)                                   |                 |          |          |                |              |
| Calcarine_L                                                                          | 4               | −86      | 14       | 7.08           | 650          |
| Postcentral_L                                                                        | −50             | −28      | 56       | 4.42           | 403          |
| Insula_R                                                                             | 38              | 12       | −8       | 4.27           | 510          |
| <i>t</i> Test Community Member No Trust > Community Member Trust (FDRc < .001 = 245) |                 |          |          |                |              |
| Calcarine_L                                                                          | 4               | −86      | 14       | 4.58           | 245          |
| <i>t</i> Test Friend No Trust > Friend Trust (FDRc < .001 = 106)                     |                 |          |          |                |              |
| Insula_R                                                                             | 40              | 24       | −6       | 4.98           | 841          |
| Supp_Motor_Area_R                                                                    | 8               | 20       | 60       | 4.56           | 1070         |
| Insula_L                                                                             | −36             | 18       | −4       | 4.45           | 316          |
| Thalamus_R                                                                           | 12              | −12      | 2        | 4.14           | 106          |
| Calcarine_L                                                                          | 4               | −86      | 14       | 3.96           | 106          |
| Frontal_Sup_R                                                                        | 18              | 54       | 24       | 3.81           | 162          |
| Frontal_Mid_R                                                                        | 36              | 6        | 46       | 3.80           | 139          |
| <i>t</i> Test Unknown Peer Trust > Friend Trust (FDRc < .001 = 203)                  |                 |          |          |                |              |
| Insula_R                                                                             | 40              | 22       | −4       | 3.81           | 203          |
| <i>t</i> Test Friend Trust > Community Member Trust (FDRc < .001 = 129)              |                 |          |          |                |              |
| Cingulum_Ant_L                                                                       | −6              | 44       | 0        | 4.99           | 1100         |
| Frontal_Mid_L                                                                        | −26             | 34       | 46       | 3.82           | 142          |
| Precuneus_L                                                                          | −8              | −46      | 32       | 3.68           | 129          |
| <i>t</i> Test Friend No Trust > Unknown Peer No Trust (FDRc < .001 = 133)            |                 |          |          |                |              |
| Supp_Motor_Area_R                                                                    | 10              | 22       | 60       | 5.04           | 2854         |
| Insula_R                                                                             | 30              | 16       | −14      | 4.52           | 133          |
| Precuneus_L                                                                          | −8              | −52      | 30       | 4.09           | 269          |

Results were FDR cluster corrected based on a threshold of  $p < .001$ . The other (reversed) contrasts did not result in significant effects.

\*Names were based on the AAL toolbox in SPM.





**Figure 4.** (A) Whole-brain  $F$  test for the interaction between Target and Choice. Results are displayed with a primary voxel-wise threshold of  $p < .001$  and FDR cluster correction of  $p < .050$ . (B) The right insula showed an interaction between Target and Choice. (C) The right SMA showed an interaction between Target and Choice.

activation to the unknown peer compared with the friend ( $p < .001$  for the right insula and  $p = .001$  for the right SMA) and community member ( $p = .001$  for the insula and  $p = .005$  for the right SMA). For no trust choices, findings revealed increased activity to the friend compared with the community member ( $p = .003$  for the right insula and  $p = .002$  for the right SMA) and unknown peer (both  $p < .001$ ).

#### Confirmatory ROI Analyses

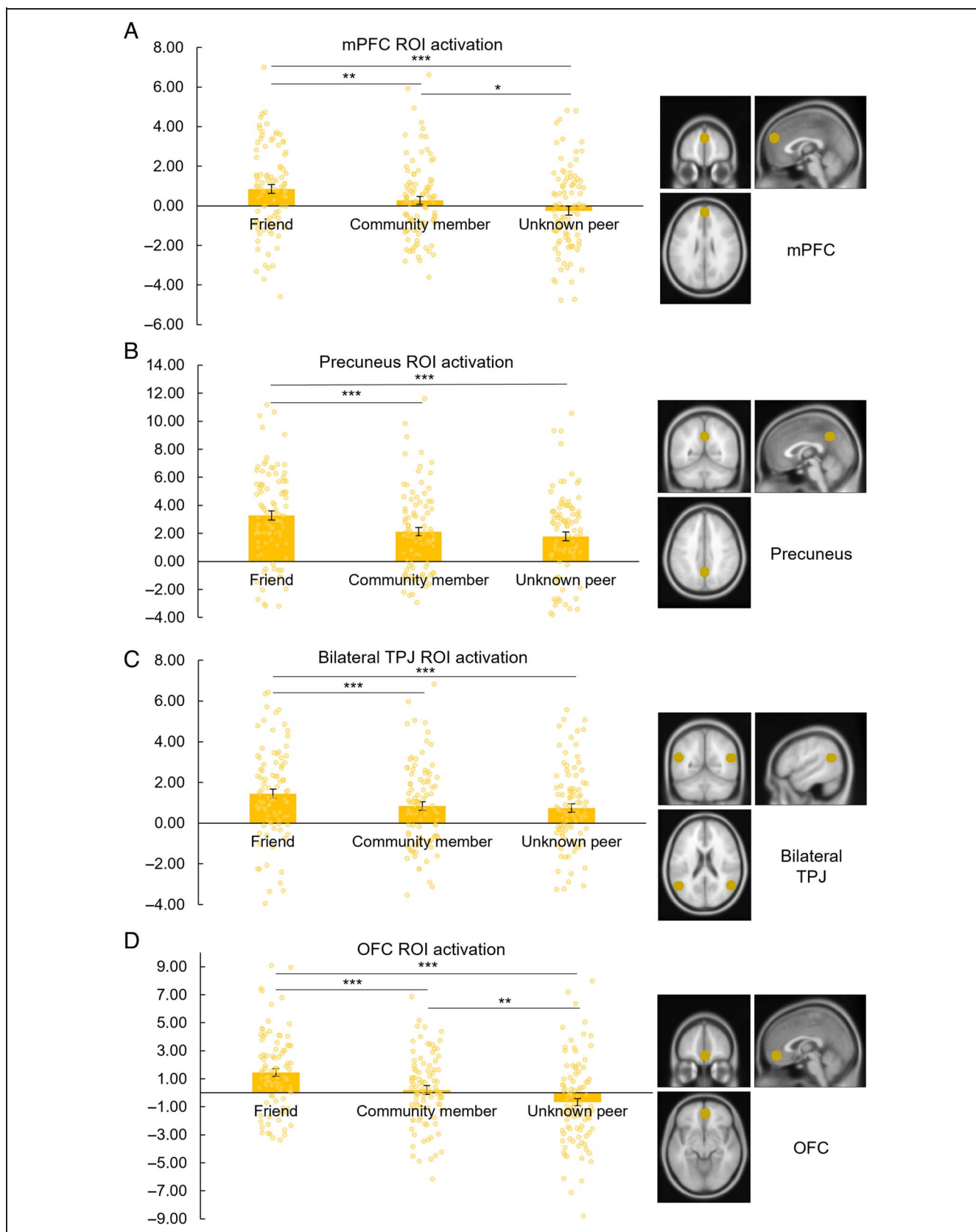
**Target only.** To examine whether ROI activity differed for the three trust targets (Hypothesis 2.1), we performed a repeated-measures ANOVA for each ROI with Target

(friend, community member, unknown peer) as within-subject variable and ROI Activity as dependent variable ( $n = 94$ ). The results were comparable for all ROIs related to mentalizing: mPFC, precuneus, and TPJ. For the mPFC, the ANOVA resulted in a main effect of Target,  $F(2, 186) = 14.47, p < .001, \eta_p^2 = .14$ , such that activation was highest for the friend ( $M = .85, SD = 2.15$ ), lower for the community member ( $M = .28, SD = 1.94$ ), and lowest for the unknown peer ( $M = -.25, SD = 2.15$ ), which is also displayed in Figure 5A. For the precuneus, there was a main effect of Target,  $F(1.85, 172.30) = 18.34, p < .001, \eta_p^2 = .17$ , such that activation was higher for the friend ( $M = 3.29, SD = 3.12$ ) compared with the community member ( $M = 2.12, SD = 2.87$ ) and unknown peer ( $M = 1.78, SD = 3.11$ ), and the latter two did not differ from each other (see Figure 5B). For the bilateral TPJ, we also detected a main effect of Target,  $F(2, 186) = 10.05, p < .001, \eta_p^2 = .10$ . Activation in the bilateral TPJ was higher for the friend ( $M = 1.44, SD = 2.15$ ) than for the community member ( $M = .85, SD = 2.08$ ) and unknown peer ( $M = .74, SD = 2.00$ ), and the latter two did not differ from each other (see Figure 5C).

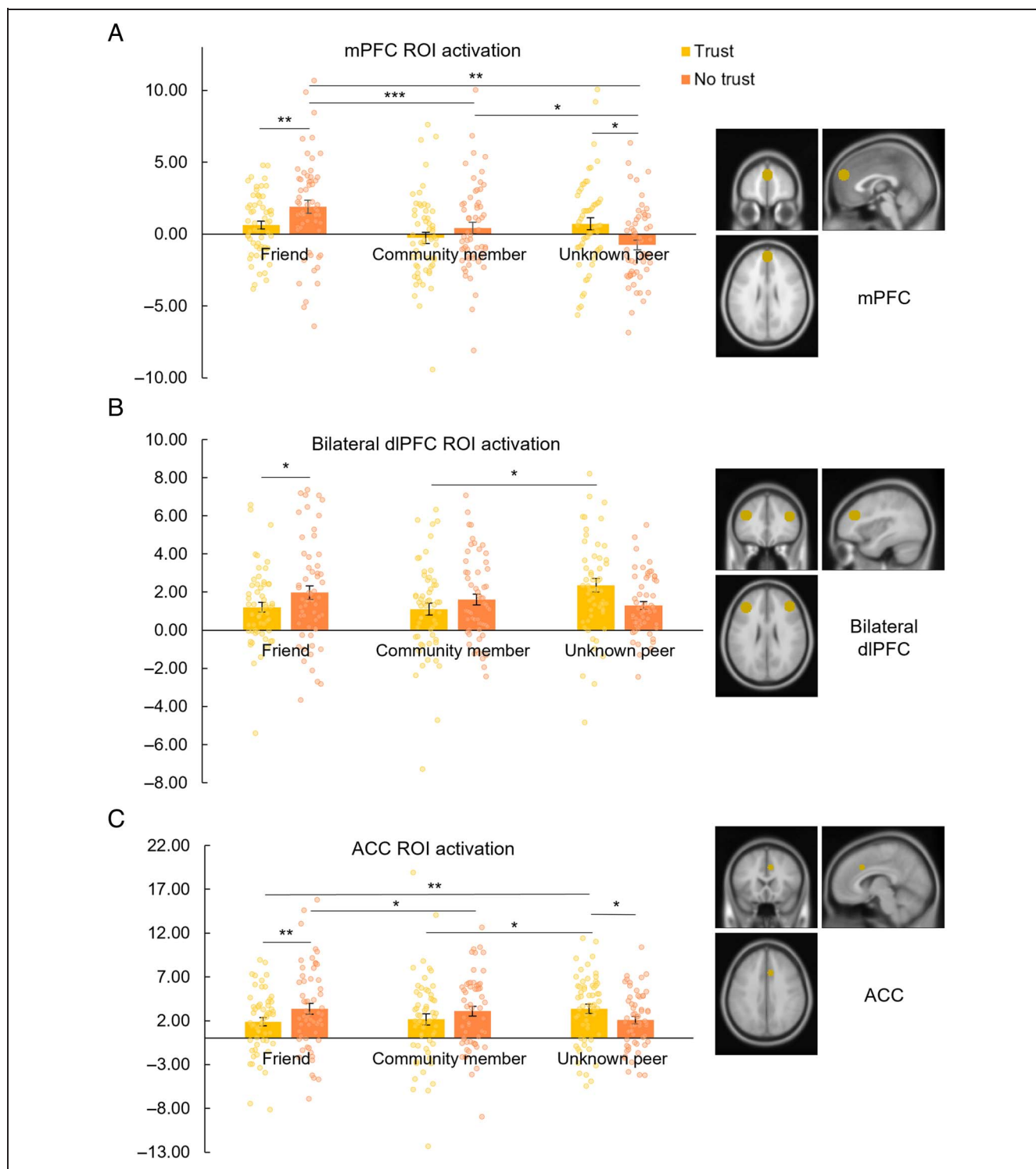
For the OFC (reward), the analysis resulted in a similar main effect of Target,  $F(2, 186) = 35.19, p < .001, \eta_p^2 = .27$ , such that activation was highest for the friend ( $M = 1.45, SD = .28$ ), lower for the community member ( $M = .19, SD = .26$ ), and lowest to the unknown peer ( $M = -.66, SD = .31$ ), which is shown in Figure 5D. The ANOVA with bilateral NAcc (reward) resulted in an uncorrected significant target effect ( $p = .036$ ), but this effect did not survive corrections for multiple testing.

Similarly, we found an uncorrected significant effect for the bilateral dlPFC (cognitive control;  $p = .032$ ), but only at an uncorrected threshold. Contrary to our hypotheses, we did not detect significant target effects for ACC.

**Target  $\times$  Choice.** To investigate whether ROI activity associated with trust behavior to different targets was dependent on the choice to trust or not to trust (Hypothesis 2.2), we performed the same analyses but included Choice (trust, no trust) as additional within-subject variable to the repeated-measures ANOVAs ( $n = 58$ ). For the mentalizing brain regions, mPFC was the only ROI that showed an interaction between Target and Choice,  $F(2, 114) = 8.86, p < .001, \eta_p^2 = .13$  (see Figure 6A). Follow-up analyses demonstrated that there were no differences between the targets for trust choices ( $p = .195$ ), but for no trust choices, activation was highest for the friend ( $M = 1.93, SD = 3.31$ ), lower for the community member ( $M = .40, SD = 3.01$ ), and lowest for the unknown peer ( $M = -.81, SD = 2.68$ ),  $F(1.67, 99.34) = 17.66, p < .001, \eta_p^2 = .23$ . In addition, follow-up analyses for each target separately showed that there was a difference in activation between trust and no trust choices for the



**Figure 5.** Task condition effects in predefined mPFC, precuneus, bilateral TPJ, and OFC. (A) The mPFC showed most activation during trust choices to the friend, more to the community member, and least to the unknown peer. (B) The precuneus showed more activation to the friend compared to the unknown peer and community member. (C) The bilateral TPJ showed more activation to the friend than unknown peer and community member. (D) The OFC showed most activation to the friend, less to the community member, and least to the unknown peer.  $*p \leq .050$ ,  $**p \leq .010$ ,  $***p \leq .001$ .



**Figure 6.** Task condition effects in predefined mPFC, dlPFC, and ACC. (A) The mPFC showed an interaction between target and choice, such that there was a difference between targets for no trust choices but not for trust choices. (B) The dlPFC showed an interaction between target and choice, such that there was a difference between the unknown peer and community member during trust choices but not during no trust choices. (C) The ACC showed an interaction between target and choice. During trust choices, activation was higher for the unknown peer compared with the community member and friend. During no trust choices, activation was higher for the friend than for the unknown peer. \* $p \leq .050$ , \*\* $p \leq .010$ , \*\*\* $p \leq .001$ .

friend,  $F(1, 67) = 7.71$ ,  $p = .007$ ,  $\eta_p^2 = .10$ ,  $M_{\text{trust}} = .65$ ,  $SD_{\text{trust}} = 2.34$ ,  $M_{\text{no trust}} = 1.90$ ,  $SD_{\text{no trust}} = 3.31$ , and unknown peer,  $F(1, 80) = 4.57$ ,  $p = .036$ ,  $\eta_p^2 = .05$ ,  $M_{\text{trust}} = .52$ ,  $SD_{\text{trust}} = 2.87$ ,  $M_{\text{no trust}} = -.37$ ,  $SD_{\text{no trust}} = 3.10$ , but

not for the community member ( $p = .283$ ). The main effect of choice was not significant ( $p = .585$ ).

Although we found no target effects in the cognitive control brain regions for bilateral dlPFC and ACC, ANOVAs

revealed that these ROIs were sensitive to the interaction between Target and Choice. The analysis for bilateral dlPFC resulted in a significant interaction between Target and Choice,  $F(2, 114) = 6.61, p = .002, \eta_p^2 = .10$ . Follow-up analyses (see also Figure 6B) showed differences in activation between targets for trust choices,  $F(1.64, 143.89) = 4.75, p = .015, \eta_p^2 = .05$ , such that bilateral dlPFC activation for trust choices was higher for the unknown peer ( $M = 1.85, SD = .29$ ) compared with the friend ( $M = 1.13, SD = .19$ ) and community member ( $M = .95, SD = .23$ ). There were no differences between targets for no trust choices ( $p = .076$ ). Follow-up comparisons per target showed that bilateral dlPFC activation between trust and no trust choices differed for the friend,  $F(1, 67) = 4.11, p = .047, \eta_p^2 = .06$ ,  $M_{\text{trust}} = 1.14, SD_{\text{trust}} = 1.85, M_{\text{no trust}} = 1.88, SD_{\text{no trust}} = 2.70$ , but not for the unknown peer ( $p = .122$ ) and community member ( $p = .191$ ). The main effect of Choice was not significant ( $p = .719$ ).

For ACC, we also observed an interaction between Target and Choice,  $F(2, 114) = 5.97, p = .003, \eta_p^2 = .10$  (see Figure 6C). Follow-up analyses showed that there were differences between targets for both trust,  $F(2, 176) = 6.60, p = .002, \eta_p^2 = .07$ , and no trust choices,  $F(1.81, 108.39) = 4.08, p = .023, \eta_p^2 = .06$ . For trust, ACC activation was higher for the unknown peer ( $M = 2.92, SD = 3.69$ ) than for the community member ( $M = 1.76, SD = 4.29$ ) and friend ( $M = 1.51, SD = 3.26$ ). For no trust, activation was higher for the friend ( $M = 3.55, SD = 4.73$ ) compared with the unknown peer ( $M = 1.95, SD = 3.18$ ), but these did not differ significantly from the community member ( $M = 3.00, SD = 4.21$ ). Follow-up analyses per target showed that ACC activation between trust and no trust choices differed for the friend,  $F(1, 67) = 8.29, p = .005, \eta_p^2 = .11, M_{\text{trust}} = 1.65, SD_{\text{trust}} = 3.41, M_{\text{no trust}} = 3.24, SD_{\text{no trust}} = 4.66$ , and the unknown peer,  $F(1, 80) = 4.50, p = .037, \eta_p^2 = .05, M_{\text{trust}} = 3.16, SD_{\text{trust}} = 3.79, M_{\text{no trust}} = 2.17, SD_{\text{no trust}} = 3.52$ , but not for the community member ( $p = .188$ ). The main effect of Choice was not significant ( $p = .295$ ).

Finally, the ANOVAs did not result in significant choice effects or interactions between target and choice for the precuneus and bilateral TPJ (mentalizing) and for the OFC and bilateral NAcc (reward).

### Confirmatory ROI Analyses with Trust Behavior

To examine whether ROI activity was associated with general trust choices (i.e., trust percentages averaged across targets; Hypothesis 2.3), we performed a repeated-measures ANOVA for each ROI that showed significant main or interaction effects for target and choice. These ANOVAs included target, or target and choice as within-subject variables, general trust as covariate, and ROI activity as dependent variable. We first performed ANOVAs for

the target-only model to maximize the number of trials included in the analysis ( $n = 94$ ) and subsequently for the target  $\times$  choice model ( $n = 58$ ).

**Target only.** The target-only model analysis resulted in a significant interaction between Target and General Trust for the mPFC only at an uncorrected threshold,  $F(2, 184) = 3.28, p = .040, \eta_p^2 = .03$ , for the bilateral TPJ at an uncorrected threshold,  $F(2, 184) = 3.14, p = .046, \eta_p^2 = .03$ , and for the precuneus at a corrected threshold,  $F(2, 184) = 20.12, p < .001, \eta_p^2 = .11$ . All patterns showed that participants with lower general trust levels displayed stronger activation during trust deliberation to friends than to unknown peers and community members (see Figure 7).

Finally, we did not detect associations between Target and General Trust for the OFC ( $p = .929$ ).

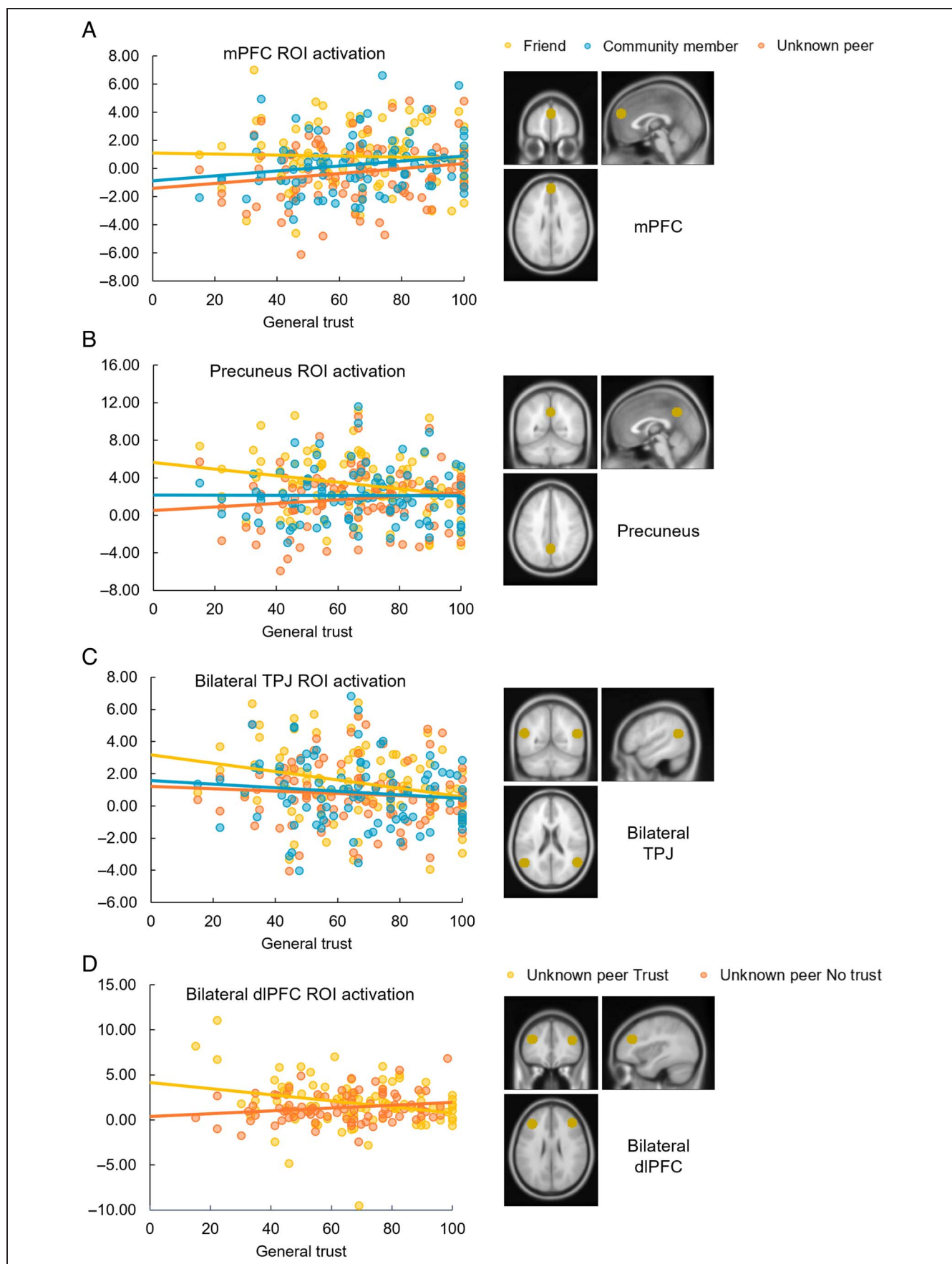
**Target  $\times$  Choice.** After adding Choice as within-subject variable to the repeated-measures ANOVA, we found an interaction for bilateral dlPFC between Choice and General Trust at an uncorrected threshold,  $F(1, 56) = 4.93, p = .031, \eta_p^2 = .08$ . To follow up on this interaction, we performed ANOVAs with Choice as within-subject variable and General Trust as covariate for each target separately. These ANOVAs showed that the interaction was only significant for the unknown peer,  $F(1, 79) = 8.70, p = .004, \eta_p^2 = .10$ , such that general trust was negatively associated with bilateral dlPFC activation during trust choices to unknown peers, and positively associated with bilateral dlPFC activation for no trust choices to unknown peers (see Figure 7D).

Finally, we found no interaction effects between Target, Choice, and General Trust for the ACC (all  $p > .05$ ).

### Confirmatory ROI Age Analyses

To investigate age effects on activation for ROIs with significant (interaction) effects of target and choice (Hypothesis 2.4), we added Linear and Quadratic age in a stepwise manner to the repeated-measures ANOVA, first for the Target-only model ( $n = 94$ ) and second for the Target  $\times$  Choice model ( $n = 58$ ). We found interactions between Linear Age and Choice for mPFC at an uncorrected threshold,  $F(1, 56) = 6.16, p = .016, \eta_p^2 = .10$ , and precuneus at a corrected threshold,  $F(1, 56) = 7.03, p = .010, \eta_p^2 = .11$ . Follow-up repeated-measures ANOVAs for each target for the mPFC revealed that the age effects were only significant for the community member,  $F(1, 67) = 4.56, p = .036, \eta_p^2 = .06$ . Compared with younger adolescents, older adolescents recruited the mPFC more strongly when not trusting the community member and less strongly during trust to the community member (see Figure 8). No post hoc tests were significant for the age effects on precuneus.

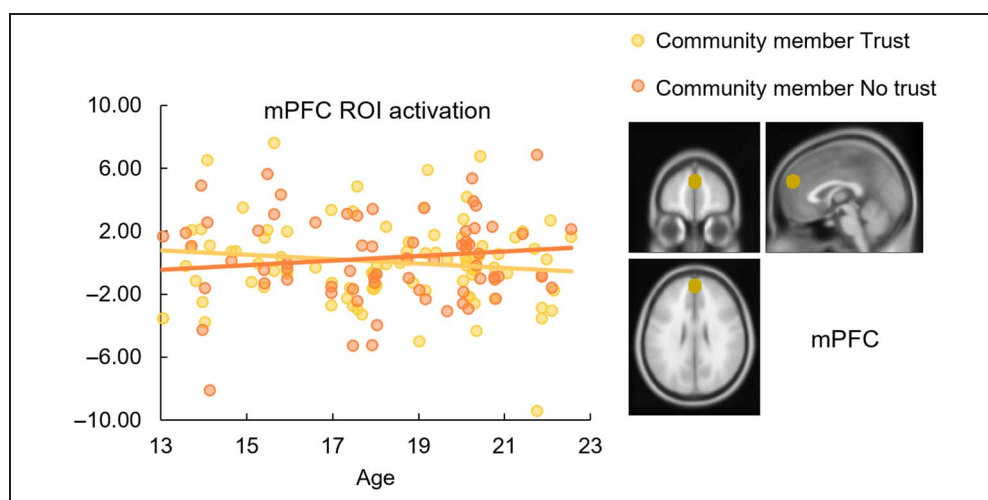




**Figure 7.** Associations between general trust and activation in the predefined (A) mPFC, (B) precuneus, (C) bilateral TPJ, and (D) bilateral dIPFC.



**Figure 8.** Association between linear age and activation in the predefined bilateral mPFC for the community member.



There were no Linear Age effects for bilateral TPJ, OFC, bilateral dlPFC, and ACC, and no Quadratic Age effects.

#### Confirmatory ROI Individual Differences Analyses

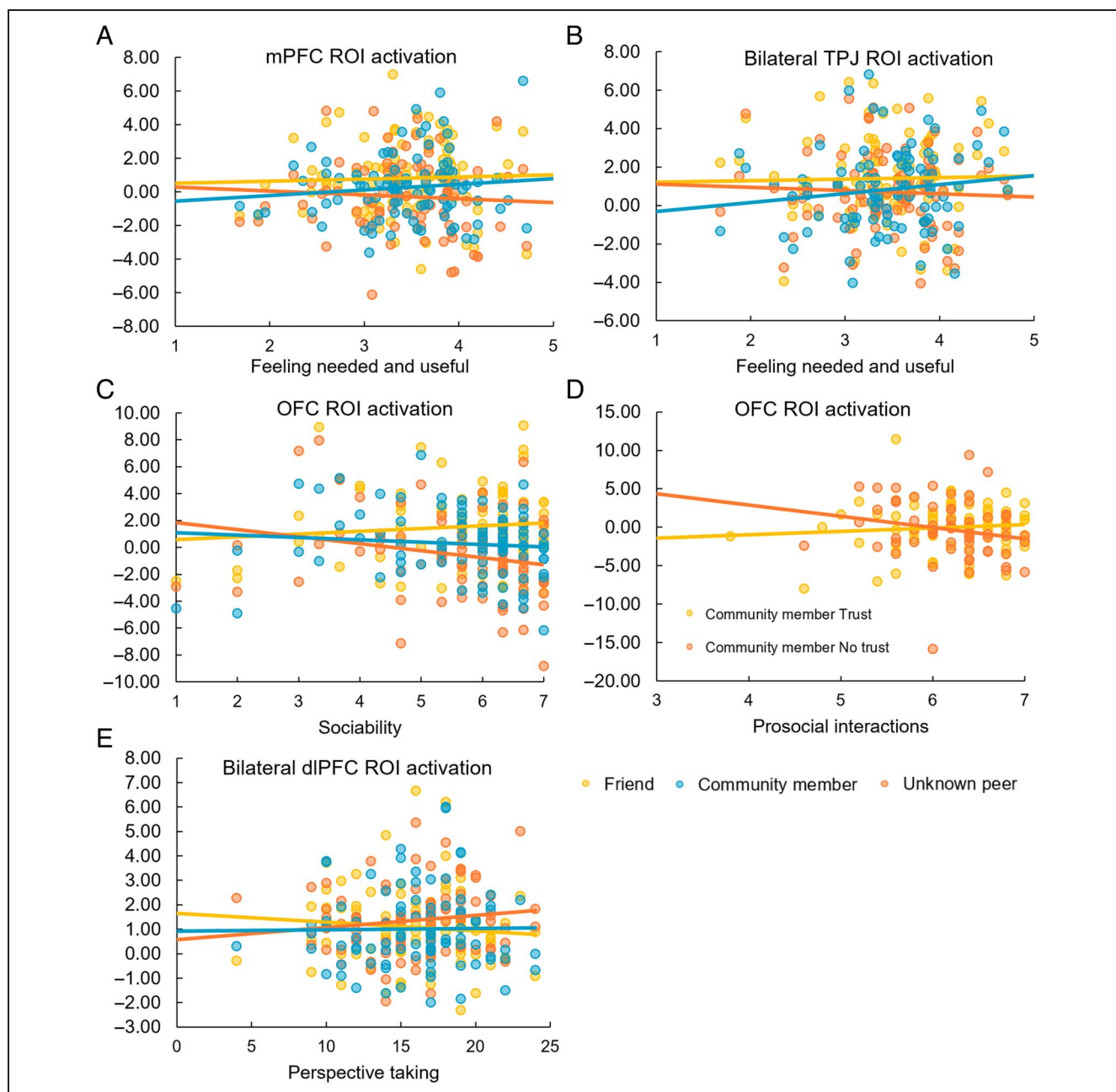
To examine whether ROI activity was associated with individual differences in perspective taking, cognitive control, sense of being needed and useful, and social reward sensitivity (Hypothesis 2.5), we added individual difference measures as covariates to the repeated-measures ANOVAs while controlling for age. For each ROI (in case of significant effects of Target and/or Target  $\times$  Choice), we performed ANOVAs for the Target-only model ( $n = 86$ ) and subsequently for the Target  $\times$  Choice model ( $n = 51$ ).

In the brain regions typically associated with mentalizing, the ROIs mPFC and bilateral TPJ showed associations with the sense of being needed and useful, only at an uncorrected threshold. For the mPFC, we found (uncorrected) significant interactions in the Target  $\times$  Choice model between target and sense of being needed and useful,  $F(2, 82) = 3.51, p = .034, \eta_p^2 = .08$ , showing that higher levels of feeling needed and useful were associated with stronger mPFC activation during trust choices to the community member than the unknown peer (see Figure 9A). Likewise, the ANOVA for bilateral TPJ resulted in an (uncorrected) interaction between target and the sense of being needed and useful,  $F(2, 152) = 3.40, p = .036, \eta_p^2 = .04$ . Follow-up analyses showed that only the difference between the unknown peer and the community member was significant,  $F(1, 76) = 7.34, p = .008, \eta_p^2 = .09$ , such that for higher levels of feeling needed and useful, adolescents recruited the bilateral TPJ more strongly during trust choices to the community member compared with the unknown peer (see Figure 9B).

For the OFC (reward), we detected an interaction between target and sociability (i.e., SRQ-A subscale) at an uncorrected threshold,  $F(2, 152) = 4.00, p = .020, \eta_p^2 = .05$ . As shown in Figure 9C, for individuals who scored higher on sociability, OFC was more strongly recruited during trust choices to the friend, compared with the unknown peer and community member. Adding choice to the ANOVA resulted in a (uncorrected) three-way interaction between Target, Choice, and Prosocial Interactions,  $F(2, 82) = 3.73, p = .028, \eta_p^2 = .08$ . To further inspect the three-way interaction, separate repeated-measures ANOVAs for each target showed that the interaction between Choice and Prosocial Interactions was only significant for the community member,  $F(1, 51) = 4.89, p = .032, \eta_p^2 = .09$  ( $p_{\text{friend}} = .197$ ;  $p_{\text{unknown peer}} = .430$ ). As displayed in Figure 9D, individuals who scored higher on prosocial interactions showed more OFC activation when trusting a community member and less activation when not trusting a community member.

For the bilateral dlPFC, the ANOVA resulted in an uncorrected interaction between Target and Perspective Taking,  $F(2, 152) = 3.56, p = .031, \eta_p^2 = .05$ . Compared with the friend and community member, adolescents with higher levels of perspective taking showed stronger recruitment of bilateral dlPFC during trust choices to the unknown peer (see Figure 9E). There was also a Target and the Sense of Feeling Needed and Useful interaction for dlPFC,  $F(2, 82) = 4.99, p = .009, \eta_p^2 = .11$ . Inspection of parameter estimates showed a negative association between feeling needed/useful and dlPFC activation for the unknown peer ( $B = -.56$ ), but a positive association for the friend ( $B = .18$ ) and community member ( $B = .21$ ), although it should be noted that these separate trends did not survive post hoc testing.

Finally, we found no associations between precuneus and ACC activation and individual differences measures.



**Figure 9.** Association between the sense of feeling needed and useful and activation in the predefined (A) mPFC and (B) bilateral TPJ. (C) Association between sociability and activation in the predefined OFC. (D) Association between prosocial interactions and activation in the predefined OFC for community members. (E) Association between perspective taking and activation in the predefined bilateral dlPFC.

## DISCUSSION

Although previous studies demonstrated that adolescents showed higher levels of trust to friends than to unknown others (Güroğlu et al., 2014), less is known about the development of trust to societal targets, such as members of a community organization. Given that a key developmental challenge of adolescence is to develop into a contributing member of society (Fuligni, 2019), this study investigated the underlying neural mechanisms of adolescents' trust to not only close (e.g., friends) and unknown

(e.g., unknown peers) targets but also to societal targets (e.g., community members) by including multiple targets to existing versions of the Trust Game. The analyses resulted in four main results. First, consistent with our previous behavioral study (Sweijen et al., 2023), adolescents showed most trust to their friend, less trust to the community member, and the least trust to the unknown peer. Moreover, we found an association between trust and feelings of being needed and useful, such that adolescents with more feelings of being needed and useful differentiated more between friends and

unknown peers in trust choices. Second, regarding differential patterns of neural activity underlying trust to close, societal, and unknown targets, we found increased activation for closer targets in the mentalizing regions mPFC, precuneus, and TPJ, and in the reward region OFC. Furthermore, whereas the cognitive control regions dlPFC and ACC showed no differences between targets per se, neural activity in these regions was dependent on the interaction between target and choice (i.e., trust or no trust). Third, older adolescents showed increased activity in the mPFC during trust to the community member, compared with no trust. Fourth, adolescents with higher feelings of being needed and useful recruited the mPFC and TPJ more during trust to the community member, compared with the unknown peer, and adolescents who score higher on prosocial interactions recruited the OFC more during trust to the community member. Finally, we also found increased activity in the dlPFC during trust to the unknown peer, compared with the friend and community member, for adolescents who show higher levels of perspective taking.

### **Adolescents' Trust to Friends, Community Members, and Unknown Peers**

One of our study aims was to examine differential patterns in adolescents' trust choices to close, societal, and unknown targets. We replicated findings of our behavioral study (Sweijen et al., 2023), indicating that adolescents showed most trust to friends, less trust to community members, and least trust to unknown peers. These findings add to the literature on adolescent trust, which previously focused on (un)familiar and (un)trustworthy others (Fett, Shergill et al., 2014; Güroğlu et al., 2014), such that we demonstrate that adolescents show trust not only to close and unknown targets but also to societal targets with whom they do not have a direct interpersonal relationship. Although a societal target is more distant in an adolescent's social world, it is important for adolescents to develop trust behavior to individuals outside of the family and friend contexts given that adolescence is characterized by developing mature social relationships and societal contributions (Crone & Fuligni, 2020; Fuligni, 2019). Future studies may investigate the extent to which trust to more distant but personally relevant targets helps adolescents to obtain these developmental goals (e.g., social competence and developing into a contributing member of society), preferably within a longitudinal design to disentangle developmental trajectories, and whether this differs from other types of prosocial behaviors (e.g., sharing and giving). Although beyond the scope of the current study, another interesting direction for future research would be to investigate the responses to trust by different targets, for example, the neural underpinnings of reciprocity percentages when receiving trust from friends relative to unknown peers.

The next aim of our study was to explore whether individual differences in advancing developmental processes throughout adolescence could explain variance in adolescent's trust choices to the different targets. Our findings demonstrated an association between trust and feeling needed and useful in multiple daily life domains (e.g., friends, family, and society). Specifically, we found that adolescents with higher feelings of being needed and useful differentiated more between their friend and an unknown peer during trust choices, which was mostly driven by a decrease in trust to an unknown peer. This finding might suggest that adolescents with higher scores on feeling needed and useful identify more with ingroup members, potentially strengthening ingroup favoritism and reducing trust toward unknown peers. This proneness to ingroup versus outgroup differentiation may be evolutionary adaptive given the benefits of belonging to an ingroup (e.g., Baumeister & Leary, 1995). This study adds to previous research showing that feeling needed and useful predicts psychological well-being (Fuligni et al., 2022a), extending it to adolescents' trust and social relationships. Future studies are needed to replicate the extent to which feeling needed and useful affects adolescents' prosocial behavior to both in- and out-group members.

We could not confirm our hypothesis that individual differences in trust were associated with behavioral indices of perspective taking, cognitive control, and social reward sensitivity, which contrasts findings of previous studies (e.g., Fett, Shergill et al., 2014; Güroğlu et al., 2014). Possibly, these relations are specific to multiround or interactive Trust Games, which are expected to require more perspective taking (Burke et al., 2020).

### **Neural Mechanisms Underlying Trust to Different Targets**

The next study aim was to investigate the underlying neural mechanisms of trust to different targets. Our findings show that trust was associated with activation in brain regions typically implicated during mentalizing, reward, and cognitive control. First, we detected most activity in the OFC and mPFC regions during trust for the friend, less for the community member, and the least for the unknown peer, which may suggest that these medial frontal midlines scale with closeness of the target. Although the mPFC and OFC show a similar pattern, the OFC may be particularly related to the reward component of trust because the increased OFC activity to closer targets may reflect increased sensitivity to social signals (Lemmers-Jansen et al., 2019; Fett, Gromann et al., 2014), which also explains the associations we found between the OFC activity and social reward sensitivity. For the mPFC, increased activity for closer targets, which was most pronounced during no trust choices, may reflect increased mentalizing (Cutler & Campbell-Meiklejohn, 2019; van den Bos et al., 2009). These effects are in line with the findings of van den Bos and colleagues (2009) who also demonstrated that the

mPFC was sensitive to the type of choice. In addition, we found an association between mPFC activity and feelings of being needed/useful. Because of its connection with several important social brain regions, the mPFC is involved in the intertwined development of self- and other-related processing, making mPFC a strong candidate for societal contributions (Crone & Fuligni, 2020), which may be supported by this association between the mPFC and feelings of being needed/useful.

Second, we found increased activity in the posterior mentalizing regions precuneus and TPJ specifically for the friend, compared with the community member and unknown peer, which is in line with previous studies showing that prosocial behavior to friends involved higher activity in mentalizing regions (van de Groep et al., 2020, 2022; Schreuders et al., 2018). These findings suggest that mentalizing regions are possibly involved in the attention shift from the self to the other (Lamm, Batson, & Decety, 2007), particularly to targets who are viewed as more similar or close to oneself such as a friend (Telzer et al., 2011). We also found that adolescents with higher levels of feeling needed and useful recruited the TPJ stronger during trust choices to the community member compared with the unknown peer, possibly reflecting increased mentalizing to a societal target. Moreover, we detected brain-behavior associations for the three mentalizing regions, such that lower general trust was associated with increased activity in the mPFC, precuneus, and TPJ, potentially reflecting increased mentalizing in adolescents who are generally less trusting to others. Taken together, our findings suggest that a key role is reserved for brain regions related to mentalizing during trust, given that we found increased activity in all brain regions previously associated with mentalizing (Blakemore & Mills, 2014; Frith & Frith, 2012; Lieberman, 2007).

Third, we found that the cognitive control regions dlPFC and ACC became increasingly responsive during trust to the unknown peer and during no trust to the friend. These findings may be explained by these choices being more unexpected and salient, as such requiring more cognitive control. Indeed, increased activity in the dlPFC and ACC has often been linked with regulating impulses (Declerck et al., 2013), which has also been demonstrated by previous studies on the role of cognitive control during trust (Lemmers-Jansen et al., 2017; Fett, Gromann et al., 2014; van den Bos et al., 2009). Furthermore, consistent with Hughes, Ambady, and Zaki (2017), higher levels of perspective taking were associated with more activity in the dlPFC when giving trust to unknown peers, suggesting that the dlPFC may play a critical role in displaying trust to outgroup members. Moreover, we also found that general trust was associated with dlPFC activity for the unknown peer, such that average levels of trust were negatively associated with activity during trust and positively associated with activity during no trust to the unknown peer. This finding contrasts that of Hughes and colleagues (2017) who found that increased activity in cognitive control regions

increases trust to outgroup members (e.g., unknown peer). These contrasting findings may be explained by differences in experimental paradigms, but future studies are needed to elucidate the role of cognitive control during trust.

Combined, our findings show that similar neural mechanisms related to mentalizing, reward, and cognitive control underly adolescents' trust to close, societal, and unknown targets. Moreover, we found that individual differences in developmental processes explain differential patterns in neural activity to close versus unknown targets. However, several results were only significant at an uncorrected threshold, particularly results regarding individual differences, so future research is needed to further investigate individual differences in neural mechanisms underlying trust.

### Developmental Effects on Trust

Many prior studies reported adolescent specific transitions in brain regions related to mentalizing, reward, and cognitive control in social interaction tasks (Blakemore & Mills, 2014). Therefore, we were particularly interested in which regions would be associated with age-related differences. Although we found no age effects on trust on a behavioral level, older adolescents showed increased activity in the mPFC when considering trusting community members, specifically when not trusting community members (relative to trusting community members), although at an uncorrected threshold. On the basis of previous studies showing age-related increases in mPFC activity during trust, reflecting advancing mentalizing skills across adolescence (Blakemore & Mills, 2014; Fett, Gromann et al., 2014; Fett, Shergill et al., 2014; van den Bos et al., 2011), our finding may suggest that these developmental effects of trust are most pronounced during trust choices to a societal partner with whom adolescents do not necessarily have an interpersonal relationship with. Previous studies showed increased activity in mentalizing regions during prosocial choices to more similar or close others, such as family members and liked others (Schreuders et al., 2018; Telzer et al., 2011; Güroğlu et al., 2008). These prior studies also showed increasing ingroup-outgroup differentiation in adolescence; therefore, the effects observed in this study may indicate that older adolescents engage in mentalizing more when not trusting a community member, but these findings should be tested in more detail in future studies. Although previous studies also detected age-related increases in ACC activity during trust, reflecting a shift from negative to positive trustworthiness expectations (Fett, Gromann et al., 2014), we did not find these age effects in cognitive control regions. Therefore, future studies are needed using longitudinal designs to elucidate how developmental effects in mentalizing are related to several types of trust and prosocial behavior to different targets.



## Strengths and Limitations

This is the first study to systemically examine underlying neural mechanisms of adolescents' trust to not only close and unknown targets but also societal targets. However, this study had several limitations that need to be addressed in future studies. First, because our study used a cross-sectional design, future studies should employ longitudinal within-subject design to disentangle developmental trajectories of trust. Second, future studies may aim to recruit more diverse samples (based on gender and ethnicity) than the current study to test whether our results can be replicated in underrepresented adolescents. Third, the study included participants aged 12 years and older where previous studies reported large changes in trust when transitioning into adolescence (Güroğlu et al., 2014). A final limitation is that data were collected during the COVID-19 pandemic, which may have impacted adolescents' trust to a societal target given that, in this period, societal issues such as trust may be under pressure.

## Conclusion

Taken together, our results show that target differentiation in adolescents' trust behavior is associated with activity in social brain regions implicated during mentalizing, reward processing, and cognitive control. Our study adds to the literature showing that adolescent trust is highly dependent on social contextual factors (i.e., the other person to whom trust is shown), not only on a behavioral but also on a neural level. Furthermore, our findings suggest that adolescents' trust behavior is mainly driven by the maturation of brain regions underlying sociocognitive processes, whereas age effects were specific to mPFC activity. Given that adolescence is characterized by the expansion of the social world and the exploration of adolescents' position within society, our study contributes to the literature on target differentiation stressing the unique role of a societal target during adolescence.

## Acknowledgments

We are thankful for all adolescents who have participated in this study. We are grateful to Jochem Spaans for his help with programming the experiment. We would also like to thank all research assistants and students who helped with the data collection: Iris Langereis, Lisanne Geurts, Christin Kuehner, Antonia Rulitschka, Lara Dilger, Daphne Jansen, Sterre van Haeringen, and Liesbeth Opschoor.

Reprint requests should be sent to Sophie W. Sweijen, Erasmus School of Social and Behavioral Sciences, Erasmus University Rotterdam, Burgemeester Oudlaan 50, 3062 PA Rotterdam, The Netherlands, or via e-mail: [sweijen@essb.eur.nl](mailto:sweijen@essb.eur.nl).

## Data Availability Statement

Research data are uploaded to Erasmus University Rotterdam's Data Repository upon acceptance (<https://doi.org/10.25397/eur.c.6474157>). Data access can be requested from the study's authors.

## Author Contributions

Sophie W. Sweijen: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Visualization; Writing—Original draft; Writing—Review & editing. Suzanne van de Groep: Conceptualization; Methodology; Writing—Original draft; Writing—Review & editing. Lysanne W. te Brinke: Conceptualization; Methodology; Writing—Original draft; Writing—Review & editing. Andrew J. Fuligni: Conceptualization; Writing—Review & editing. Eveline A. Crone: Conceptualization; Funding acquisition; Writing—Original draft; Writing—Review & editing.

## Funding Information

This work was supported by an innovative ideas grant of the European Research Council (ERC CoG PROSOCIAL 681632 to E. A. C.).

## Diversity in Citation Practices

Retrospective analysis of the citations in every article published in this journal from 2010 to 2021 reveals a persistent pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the *Journal of Cognitive Neuroscience (JoCN)* during this period were  $M(\text{an})/M = .407$ ,  $W(\text{oman})/M = .32$ ,  $M/W = .115$ , and  $W/W = .159$ , the comparable proportions for the articles that these authorship teams cited were  $M/M = .549$ ,  $W/M = .257$ ,  $M/W = .109$ , and  $W/W = .085$  (Postle and Fulvio, *JoCN*, 34:1, pp. 1–3). Consequently, *JoCN* encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article's gender citation balance. The authors of this paper report its proportions of citations by gender category to be:  $M/M = .31$ ;  $W/M = .15$ ;  $M/W = .15$ ;  $W/W = .39$ .

## REFERENCES

- Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117, 497–529. <https://doi.org/10.1037/0033-2909.117.3.497>, PubMed: 7777651
- Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10, 122–142. <https://doi.org/10.1006/game.1995.1027>
- Blakemore, S. J. (2008). The social brain in adolescence. *Nature Reviews Neuroscience*, 9, 267–277. <https://doi.org/10.1038/nrn2353>, PubMed: 18354399
- Blakemore, S. J., & Mills, K. L. (2014). Is adolescence a sensitive period for sociocultural processing? *Annual Review of Psychology*, 65, 187–207. <https://doi.org/10.1146/annurev-psych-010213-115202>, PubMed: 24016274
- Brett, M., Anton, J. L., Valabregue, R., & Poline, J. B. (2002). Region of interest analysis using an SPM toolbox.

- Neuroimage*, 16, 497. [https://doi.org/10.1016/S1053-8119\(02\)90013-3](https://doi.org/10.1016/S1053-8119(02)90013-3)
- Burke, S. M., van de Groep, S., Brandner, P., & Crone, E. A. (2020). Neurocognitive developmental changes in trust and reciprocity across adolescence. In K. C. Kadosh (Ed.), *The Oxford handbook of developmental cognitive neuroscience* (pp. 1–24). Oxford Academic. <https://doi.org/10.1093/oxfordhb/9780198827474.013.26>
- Crone, E. A., & Dahl, R. E. (2012). Understanding adolescence as a period of social-affective engagement and goal flexibility. *Nature Reviews Neuroscience*, 13, 636–650. <https://doi.org/10.1038/nrn3313>, PubMed: 22903221
- Crone, E. A., & Fuligni, A. J. (2020). Self and others in adolescence. *Annual Review of Psychology*, 71, 447–469. <https://doi.org/10.1146/annurev-psych-010419-050937>, PubMed: 31337274
- Crone, E. A., Sweijen, S. W., te Brinke, L. W., & van de Groep, S. (2022). Pathways for engaging in prosocial behavior in adolescence. *Advances in Child Development and Behavior*, 63, 149–190. <https://doi.org/10.1016/bs.acdb.2022.03.003>, PubMed: 35871821
- Cutler, J., & Campbell-Meiklejohn, D. (2019). A comparative fMRI meta-analysis of altruistic and strategic decisions to give. *Neuroimage*, 184, 227–241. <https://doi.org/10.1016/j.neuroimage.2018.09.009>, PubMed: 30195947
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, 44, 113–126. <https://doi.org/10.1037/0022-3514.44.1.113>
- Declerck, C. H., Boone, C., & Emonds, G. (2013). When do people cooperate? The neuroeconomics of prosocial decision making. *Brain and Cognition*, 81, 95–117. <https://doi.org/10.1016/j.bandc.2012.09.009>, PubMed: 23174433
- Delgado, M. R., Frank, R. H., & Phelps, E. A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nature Neuroscience*, 8, 1611–1618. <https://doi.org/10.1038/nn1575>, PubMed: 16222226
- Dumontheil, I., Apperly, I. A., & Blakemore, S. J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental Science*, 13, 331–338. <https://doi.org/10.1111/j.1467-7687.2009.00888.x>, PubMed: 20136929
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G\*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175–191. <https://doi.org/10.3758/bf03193146>, PubMed: 17695343
- Fett, A. K. J., Gromann, P. M., Giampietro, V., Shergill, S. S., & Krabbendam, L. (2014). Default distrust? An fMRI investigation of the neural development of trust and cooperation. *Social Cognitive and Affective Neuroscience*, 9, 395–402. <https://doi.org/10.1093/scan/nss144>, PubMed: 23202661
- Fett, A. K. J., Shergill, S. S., Gromann, P. M., Dumontheil, I., Blakemore, S. J., Yakub, F., et al. (2014). Trust and social reciprocity in adolescence—A matter of perspective-taking. *Journal of Adolescence*, 37, 175–184. <https://doi.org/10.1016/j.adolescence.2013.11.011>, PubMed: 24439623
- Foulkes, L., Neumann, C. S., Roberts, R., McCrory, E., & Viding, E. (2017). Social Reward Questionnaire—Adolescent version and its association with callous–unemotional traits. *Royal Society Open Science*, 4, 160991. <https://doi.org/10.1098/rsos.160991>, PubMed: 28484617
- Frith, C. D., & Frith, U. (2012). Mechanisms of social cognition. *Annual Review of Psychology*, 63, 287–313. <https://doi.org/10.1146/annurev-psych-120710-100449>, PubMed: 21838544
- Fuligni, A. J. (2019). The need to contribute during adolescence. *Perspectives on Psychological Science*, 14, 331–343. <https://doi.org/10.1177/1745691618805437>, PubMed: 30562473
- Fuligni, A. J. (2020). Is there inequality in what adolescents can give as well as receive? *Current Directions in Psychological Science*, 29, 405–411. <https://doi.org/10.1177/0963721420917738>, PubMed: 33758477
- Fuligni, A. J., Smola, X. A., & Al Salek, S. (2022a). Feeling needed and useful during the transition to young adulthood. *Journal of Research on Adolescence*, 32, 1259–1266. <https://doi.org/10.1111/jora.12680>, PubMed: 34545640
- Fuligni, A. J., Smola, X. A., & Al Salek, S. (2022b). Youth's experiences of the devaluing of their contributions due to their ethnicity and gender. *Journal of Youth and Adolescence*, 51, 1667–1678. <https://doi.org/10.1007/s10964-022-01617-1>, PubMed: 35508824
- Güroğlu, B. (2021). Adolescent brain in a social world: Unravelling the positive power of peers from a neurobehavioral perspective. *European Journal of Developmental Psychology*, 18, 471–493. <https://doi.org/10.1080/17405629.2020.1813101>
- Güroğlu, B., van den Bos, W., & Crone, E. A. (2014). Sharing and giving across adolescence: An experimental study examining the development of prosocial behavior. *Frontiers in Psychology*, 5, 219. <https://doi.org/10.3389/fpsyg.2014.00291>, PubMed: 24782796
- Güroğlu, B., Haselager, G. J. T., van Lieshout, C. F. M., Takashima, A., Rijpkema, M., & Fernandez, G. (2008). Why are friends special? Implementing a social interaction simulation task to probe the neural correlates of friendship. *Neuroimage*, 39, 903–910. <https://doi.org/10.1016/j.neuroimage.2007.09.007>, PubMed: 17964185
- Hughes, B. L., Ambady, N., & Zaki, J. (2017). Trusting outgroup, but not ingroup members, requires control: Neural and behavioral evidence. *Social Cognitive and Affective Neuroscience*, 12, 372–381. <https://doi.org/10.1093/scan/nsw139>, PubMed: 27798248
- Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology*, 32, 865–889. <https://doi.org/10.1016/j.joep.2011.05.007>
- Karan, M., Lazar, L., Leschak, C. J., Galvan, A., Eisenberger, N. I., Uy, J. P., et al. (2022). Giving to others and neural processing during adolescence. *Developmental Cognitive Neuroscience*, 56, 101128. <https://doi.org/10.1016/j.dcn.2022.101128>, PubMed: 35759828
- King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., & Montague, P. R. (2005). Getting to know you: Reputation and trust in a two-person economic exchange. *Science*, 308, 78–83. <https://doi.org/10.1126/science.1108062>, PubMed: 15802598
- Konovalov, A., Hill, C., Daunizeau, J., & Ruff, C. C. (2021). Dissecting functional contributions of the social brain to strategic behavior. *Neuron*, 109, 3323–3337. <https://doi.org/10.1016/j.neuron.2021.07.025>, PubMed: 34407389
- Lahno, B. (1995). Trust, reputation, and exit in exchange relationships. *Journal of Conflict Resolution*, 39, 495–510. <https://doi.org/10.1177/0022002795039003005>
- Lamm, C., Batson, C. D., & Decety, J. (2007). The neural substrate of human empathy: Effects of perspective-taking and cognitive appraisal. *Journal of Cognitive Neuroscience*, 19, 42–58. <https://doi.org/10.1162/jocn.2007.19.1.42>, PubMed: 17214562
- Lemmers-Jansen, I. L. J., Fett, A. K. J., Shergill, S. S., van Kesteren, M. T. R., & Krabbendam, L. (2019). Girls–boys: An investigation of gender differences in the behavioral and neural mechanisms of trust and reciprocity in adolescence. *Frontiers in Human Neuroscience*, 13, 257. <https://doi.org/10.3389/fnhum.2019.00257>, PubMed: 31427936
- Lemmers-Jansen, I. L. J., Krabbendam, L., Veltman, D. J., & Fett, A. K. J. (2017). Boys vs. girls: Gender differences in the neural

- development of trust and reciprocity depend on social context. *Developmental Cognitive Neuroscience*, 25, 235–245. <https://doi.org/10.1016/j.dcn.2017.02.001>, PubMed: 28258836
- Lieberman, M. D. (2007). Social cognitive neuroscience: A review of core processes. *Annual Review of Psychology*, 58, 259–289. <https://doi.org/10.1146/annurev.psych.58.110405.085654>, PubMed: 17002553
- Lockwood, P. L., Apps, M. A. J., & Chang, S. W. C. (2020). Is there a “social” brain? Implementations and algorithms. *Trends in Cognitive Sciences*, 24, 802–813. <https://doi.org/10.1016/j.tics.2020.06.011>, PubMed: 32736965
- Miezin, F. M., Maccotta, L., Ollinger, J. M., Petersen, S. E., & Buckner, R. L. (2000). Characterizing the hemodynamic response; effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. *Neuroimage*, 11, 735–759. <https://doi.org/10.1006/nimg.2000.0568>, PubMed: 10860799
- Rilling, J. K., & Sanfey, A. G. (2011). The neuroscience of social decision-making. *Annual Review of Psychology*, 62, 23–48. <https://doi.org/10.1146/annurev.psych.121208.131647>, PubMed: 20822437
- Sankoh, A. J., Huque, M. F., & Dubey, S. D. (1997). Some comments on frequently used multiple endpoint adjustment methods in clinical trials. *Statistics in Medicine*, 16, 2529–2542. [https://doi.org/10.1002/\(SICI\)1097-0258\(19971130\)16:22<2529::AID-SIM692>3.0.CO;2-J](https://doi.org/10.1002/(SICI)1097-0258(19971130)16:22<2529::AID-SIM692>3.0.CO;2-J), PubMed: 9403954
- Schreuders, E., Klapwijk, E. T., Will, G. J., & Güroğlu, B. (2018). Friend versus foe: Neural correlates of prosocial decisions for liked and disliked peers. *Cognitive, Affective, & Behavioral Neuroscience*, 18, 127–142. <https://doi.org/10.3758/s13415-017-0557-1>, PubMed: 29318509
- Schreuders, E., Smeekens, S., Cillessen, A. H. N., & Güroğlu, B. (2019). Friends and foes: Neural correlates of prosocial decisions with peers in adolescence. *Neuropsychologia*, 129, 153–163. <https://doi.org/10.1016/j.neuropsychologia.2019.03.004>, PubMed: 30871971
- Sjitsma, H., van Buuren, M., Hollarek, M., Walsh, R., Lee, N., Braams, B., et al. (2023). Social network positions, trust behavior and its neural mechanisms in young adolescents. *Neuroimage*, 268, 119882. <https://doi.org/10.1016/j.neuroimage.2023.119882>
- Sutter, M., & Kocher, M. G. (2007). Trust and trustworthiness across different age groups. *Games and Economic Behavior*, 59, 364–382. <https://doi.org/10.1016/j.geb.2006.07.006>
- Sweijen, S. W., te Brinke, L. W., van de Groep, S., & Crone, E. A. (2023). Adolescents’ trust and reciprocity toward friends, unknown peers, and community members. *Journal of Research on Adolescence*. Advance online publication. <https://doi.org/10.1111/jora.12888>
- Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2008). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality*, 72, 271–324. <https://doi.org/10.1111/j.0022-3506.2004.00263.x>, PubMed: 15016066
- Telzer, E. H., Masten, C. L., Berkman, E. T., Lieberman, M. D., & Fuligni, A. J. (2011). Neural regions associated with self control and mentalizing are recruited during prosocial behaviors towards the family. *Neuroimage*, 58, 242–249. <https://doi.org/10.1016/j.neuroimage.2011.06.013>, PubMed: 21703352
- Uitenbroek, D. G. (1997). SISA binomial. <https://www.quantitativeskills.com/sisa/calculations/bonfer.htm>
- van de Groep, S., Meuwese, R., Zanolie, K., Güroğlu, B., & Crone, E. A. (2018). Developmental changes and individual differences in trust and reciprocity in adolescence. *Journal of Research on Adolescence*, 30, 192–208. <https://doi.org/10.1111/jora.12459>, PubMed: 30325088
- van de Groep, S., Zanolie, K., Burke, S. M., Brandner, P., Fuligni, A. J., & Crone, E. A. (2022). Growing in generosity? The effects of giving magnitude, target, and audience on the neural signature of giving in adolescence. *Developmental Cognitive Neuroscience*, 54, 101084. <https://doi.org/10.1016/j.dcn.2022.101084>, PubMed: 35180635
- van de Groep, S., Zanolie, K., & Crone, E. A. (2020). Giving to friends, classmates, and strangers in adolescence. *Journal of Research on Adolescence*, 30, 290–297. <https://doi.org/10.1111/jora.12491>, PubMed: 30861246
- van den Bos, W., van Dijk, E., Westenberg, M., Rombouts, S. A. R. B., & Crone, E. A. (2009). What motivates repayment? Neural correlates of reciprocity in the trust game. *Social Cognitive and Affective Neuroscience*, 4, 294–304. <https://doi.org/10.1093/scan/nsp009>, PubMed: 19304843
- van den Bos, W., van Dijk, E., Westenberg, M., Rombouts, S. A. R. B., & Crone, E. A. (2011). Changing brains, changing perspectives: The neurocognitive development of reciprocity. *Psychological Science*, 22, 60–70. <https://doi.org/10.1177/0956797610391102>, PubMed: 21164174
- van den Bos, W., Westenberg, M., van Dijk, E., & Crone, E. A. (2010). Development of trust and reciprocity in adolescence. *Cognitive Development*, 25, 90–102. <https://doi.org/10.1016/j.cogdev.2009.07.004>
- van Hoorn, J., van Dijk, E., Güroğlu, B., & Crone, E. A. (2016). Neural correlates of prosocial peer influence on public goods game donations during adolescence. *Social Cognitive and Affective Neuroscience*, 11, 923–933. <https://doi.org/10.1093/scan/nsw013>, PubMed: 26865424